

# Behavioural Pattern Identification and Prediction in Intelligent Environments

Sawsan Mahmoud, Ahmad Lotfi and Caroline Langensiepen

*School of Science and Technology, Nottingham Trent University  
Nottingham, NG11 8NS, United Kingdom*

*Email: {sawsan.mahmoud, ahmad.lotfi, caroline.langensiepen}@ntu.ac.uk*

---

## Abstract

In this paper, the application of soft computing techniques in prediction of an occupant's behaviour in an inhabited intelligent environment is addressed. In this research, daily activities of elderly people who live in their own homes suffering from dementia are studied. Occupancy sensors are used to extract the movement patterns of the occupant. The occupancy data is then converted into temporal sequences of activities which are eventually used to predict the occupant behaviour. To build the prediction model, different dynamic recurrent neural networks are investigated. Recurrent neural networks have shown a great ability in finding the temporal relationships of input patterns. The experimental results show that non-linear autoregressive network with exogenous inputs model correctly extracts the long term prediction patterns of the occupant and outperformed the Elman network. The results presented here are validated using data generated from a simulator and real environments.

*Keywords:* Intelligent Environments, Occupancy Monitoring, Binary Time series, Health Monitoring, Recurrent Neural Networks, NARX, Elman Network.

---

## 1. Introduction

Dealing with associated aging problems and improving the lifestyle of elderly or people of special needs in their own homes has become an important issue in the area of health monitoring. A significant health care budget saving for both government and individual families can be achieved by reducing the number of elderly people in hospitals. This can be assisted by early detection of health problems to alert and protect those people from

getting into a worse situation [1][2]. Early detection of abnormal behaviour can provide us with important information in the early stages. For instance, early detection of change in sleeping patterns or changes in room occupancy pattern would act as a preventive health monitoring system. To achieve this, the environment where the individual lives should have the capability of collecting important information about the occupant behaviour and detect gradual changes in the behaviour [3]. This environment is a space in a real world which is equipped with a large number of tools and devices such as sensors, actuators and computing components [4][5].

In the research reported here, the intelligent environment is used to monitor continuously any changes in the behavioural patterns of the occupant; mainly elderly people suffering from dementia. The advantage of using such a solution is that patients will be able to live in the comfort of their own homes and at the same time achieve relatively high degrees of confidence that they will be helped when required. Figure 1 illustrates a schematic diagram of the monitoring and interaction system architecture in an intelligent environment. The data collected from the sensor network are communicated with a base station and eventually stored in a central database. The data acquired includes the occupancy of different areas, environmental attributes, and interactions between an occupant and devices. Sensory devices are responsible for data collection and a variety of sensors are readily available to perform this task. The following list gives the detail of typical sensors:

- Passive Infra-red Sensors (PIR) or motion detectors are sensitive to the movements of living objects. PIR motion sensors respond to changes in heat in the form of infra-red radiation. They are used to identify the movement and then the movement pattern is interpreted as the occupancy.
- Door/Window entry point sensors are on/off switches which can detect the open and closed status of a door/window. Door entry point sensors are relatively reliable as they clearly represent the movement activities.
- Bed/sofa pressure sensors are used to measure the presence in and usage of these areas.

The aim of this study is to examine the ability of Recurrent Neural Networks (RNNs) in terms of modelling and prediction of sensory data to monitor the Activities of Daily Living (ADL) of an occupant living in an Inhabited Intelligent Environment (IIE). The daily behaviour patterns of the occupant are extracted using a wireless sensor network system. This

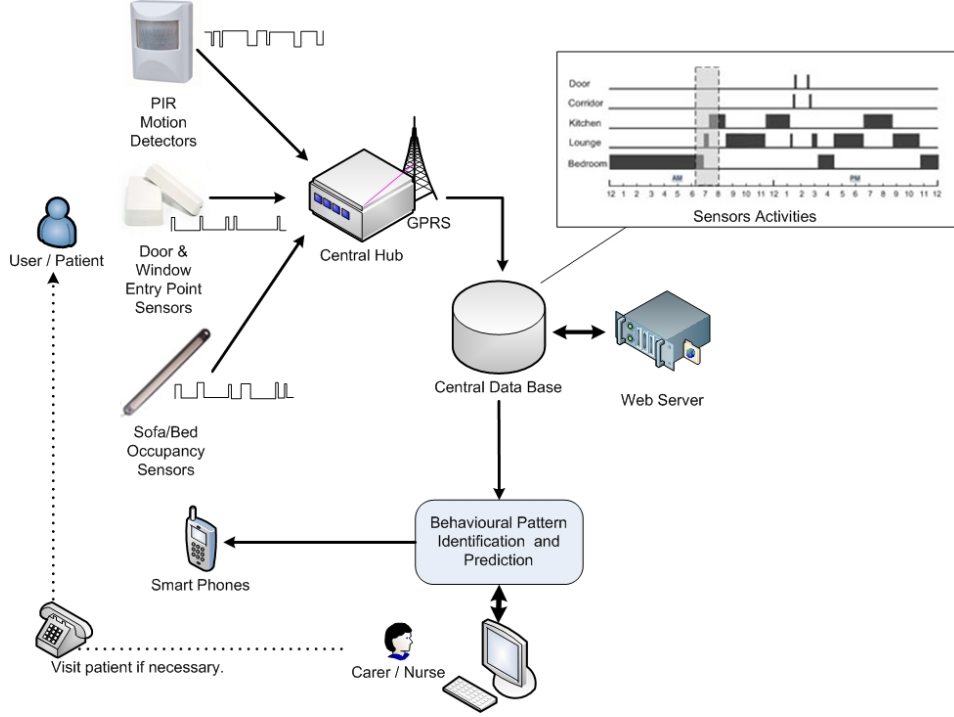


Figure 1: An overview of monitoring and interaction system architecture.

information is then used to build a behavioural model of the occupant which ultimately is used to predict the future values representing the expected occupancy and other activities. Collected information are in a binary format, and they are represented as binary time series.

Binary time series extracted from occupancy sensors are usually sparse and contain many repeated constant values. Figure 2 illustrates an intermittent binary time series generated from an occupancy sensor. A constant value (usually zero) appears many times in the series. Based on the temporal relationship of the time series, it is difficult to identify the daily movement

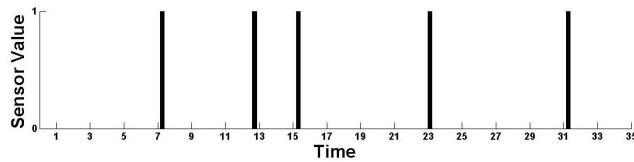


Figure 2: Intermittent binary time series extracted from an occupancy sensor.

pattern of the occupant. In our earlier work published in [6], we realised this issue since the actual sensor data were used to identify and predict the behaviour of an elderly person.

In this paper, the binary intermittent time series representing the occupancy is first converted into a compressed non-binary time series. This is achieved by calculating the start-time and stop-time for each movement from one area into another. After data conversion, the new compressed time series is used to learn and eventually to predict an occupant behaviour in an IIE. The series may be modelled and predicted as a supervised training phase. Supervised learning with an explicit training phase offers a promising approach to the activity recognition problem.

Non-linear Autoregressive network with eXogenous inputs (NARX), a recurrent dynamic network with feedback connections enclosing several layers of the network, is used to predict the future values of the series representing future activities of an occupant in an IIE. This model has feedback connections to the input layer from the output neuron which allow the output to be included in prediction of the ADL for an occupant. The key advantages of NARX network over other recurrent neural networks are its generalisation and convergence at learning long term dependencies. By long term dependencies, we mean the ability of the network to remember information that is stored for a long period of time [7].

This paper is organised as follows: Section 2 presents a brief survey of related works including any limitations. Section 3 present the pre-processing of the binary time series followed by an explanation of the prediction models used in this study for occupancy prediction in Section 4. Section 5 presents the prediction error measurements used to evaluate the models. Experimental results and data collection via two environments (simulated and real environments) are discussed in Section 6. Final conclusions are drawn in Section 7.

## 2. Related Works

A key challenge in intelligent environments research is to find a technique that can effectively model the human occupant behaviour. In this area, extensive research has been carried out to tackle the problem of activity extraction and prediction. Most of this research has been done using statistical methods. These methods are used to find the dependence and correlations between the temporal data generated from sensors and ultimately identify the behaviour of an occupant. Probabilistic models become good statistical methods to identify human behaviour as they are capable of

representing random variables, dependencies and temporal variation within data [8]. Probabilistic models could be temporal such as Hidden Markov Models (HMM) [9][10] and Bayesian belief networks [11]. HMM consists of a number of hidden states and observations, and is used to model human behaviour. However, the main problems of HMM include the lack of hierarchy in this kind of modelling and the difficulties experienced in processing large amounts low-level sensory data. Using HMM in time series predictions require a large number of time series runs from HMM as the length of time series is increased [8][10]. Bayesian belief network is a statistical method which provides a more general framework to model human behaviour. These methods are used as a tool to process uncertain and incomplete data. The major problem of Bayesian belief networks is the inflexibility of exact probabilistic inference [8][12][13][14].

Some of the above mentioned problems are addressed using soft computing techniques. Different combinations of Artificial Neural Networks (ANNs) are used in learning the occupant's activities in an IIE. For example; as reported in [12] a One Pass Neural Network is proposed to find the activities of a user. The authors have introduced a new layer to the design of the network where the outputs of neural network are inputs to this layer. This layer consists of several cells including: static cell (expert knowledge), dynamic cell (temporal order) and decision making cell. The layer helps in differentiating between normal and abnormal behaviours based on the frequencies of ADL.

Multi-Layer Perceptron (MLP) neural networks are used to identify the movements data collected from a WSN [15][16]. For instance, the authors in [17] have applied different algorithms to recognize the age categories of data representing waking pattern and to identify the change in volunteers behaviour change. These algorithms are: MLP, decision tree, support vector classifier, Naive Bayes and Bayesnet. MLP gives the highest accuracy in classifying the categories, although the size of training and testing data sets is small. In addition, in [18] a feed forward neural network with back propagation is implemented to control the basic occupant's living conditions such as air, heating, lighting, ventilation, and water heating. The authors in [19] have proposed a special kind of self organising map for clustering daily human activities. The proposed self adaptive neural network is called a growing self organising map. Much research has used ANNs in modelling and prediction of time series data. Most of the reported research has focused on using feed forward neural networks [20][21]. These networks do not have any feedback connections in their architectures. Therefore, they do not take into account the temporal dependencies between the data.

Recurrent neural networks are widely used to deal with many dynamical and non-linear problems, including time series prediction. RNNs are computationally more powerful than feed forward networks and valuable approximation result have been obtained for prediction problems [22][23]. In [24] and [25] a temporal neural-network based embedded agent is used which can work with real-time data from unobtrusive low-level sensors and actuators. An algorithm namely Selective Recursive Kernel based Learning (SRKL) is introduced in [26] and [27]. The SRKL algorithm identifies a model which has a sparse solution and an on-line identification of non-linear Multi-Input-Multi-Output with NARX system is investigated in [27]. SRKL is able to learn the process adaptively with relatively small computation load and few samples.

There are numerous researches that attempt to represent and classify the behaviour activities extracted from an IIE. For example, the authors in [28] have incorporated a sequential pattern identification method to represent the user's movements during the day. Using this approach, the activities which are performed by an individual are represented using a single continuous vertical trajectory.

Supervised learning techniques such as Support Vector Machine (SVM) also show promising identification performance. For example, SVM is used in [29] to predict the occupant behaviour. The process enables predicting householder's activities for frequent daily activities in the house such as grooming, eating, sleeping, having breakfast, etc. The activity is identified for assessment initially. Then, SVM classifier is trained, using the datasets collected from sensors where the users perform their activities, by learning the user's habit. However, the results are limited to the activities that are carried out at early morning only. Machine learning techniques are also employed in [30] to predict an inhabitant behaviour patterns, activities and common communication inside a home. The extracted data is utilised in automating decision making and improving an inhabitant's comfort, safety and efficiency. The movement of the inhabitant is predicted based on principles of information theory. An algorithm is set out on sequence matching in order to predict an inhabitant communications with the smart home, and also to identify significant patterns of inhabitant activity.

Application of fuzzy data analysis in intelligent environments have already been investigated in telecare systems [31][32]. Due to the high degree of uncertainty, it is shown that a fuzzy system would be able to answer high level queries such as "is the person eating regularly". An adaptive fuzzy control system reported in [33] has made an attempt to propose an adaptive home to adapt to its inhabitants' living style. Similar applications using

embedded agents are reported in [34] and [35].

A system for ADL recognition system is proposed in [36] using fuzzy logic. Fuzzy set theory is applied to monitor the ADL of an occupant to offer him/her a safe environment. In [37], a fuzzy predictor model is used to build the prediction model and then the results are compared with the traditional time series prediction models such as ARMA, adaptive network-based fuzzy inference system and transductive neuro-fuzzy inference model with weighted normalization. In [38], a fuzzy learning and adaptation approach for agents, called an Adaptive On-line Fuzzy Inference System, is proposed for ubiquitous computing environments. This approach consists of five phases including: monitoring the users' behaviour, capturing the actual data associated with their activities, computing the fuzzy membership functions from the input/output data, generating the fuzzy rules from the data and the agent control the learning and adaptation process. Thus, the intelligent agent has learnt, predicted and adapted to the needs of the user. To validate the proposed approach, it is compared with other computational techniques such MLP neural networks, genetic programming and the adaptive neuro fuzzy inference system. The experiments are conducted on only five days in the intelligent Dormitory (iDorm) real environment.

### 3. Binary Time Series Preprocessing

Data preprocessing is an essential step in working with any ANNs. As stated in Section 1, using the binary signal is not an efficient way of representing a large volume of intermittent data. This form of signal representation could limit the generalisation of any network for data collected over a long period of time. A binary time series represents the occupancy activities collected from the door entry sensors and movement sensors. The series contain a large sequence of 0's and 1's value since the sensor value is either 'on' or 'off' (see Figure 2). Therefore, it is essential to convert a binary series into a more suitable format to be fed as inputs to recurrent neural network models.

We have investigated and tested different forms of data representation for binary series as discussed in the following sections.

#### 3.1. High to Low Frequency Sampling

High to low frequency conversion is used when data is re-sampled at a lower rate, e.g. data recorded in seconds converted into hours; daily data into monthly data. Under these conditions, the actual time series data are accumulated over time to form a new, less frequent time series, thus helping

with visualisation and data processing. This method is based on a particular choice of statistical method namely, sum, mean, standard deviation, or any other methods which can be used to aggregate the transactions within a specific time period. This technique does not change the fact that the series is still in a sparse format.

### 3.2. Start-time and Stop-time

Start-time and stop-time method is a conversion of the series where the binary time sequence is represented as a compressed time-sampled sequence of starting time and stopping time for each activity. This form of data representation is shown to be more suitable for any modelling and prediction [11]. This form of representation is formally defined in the remaining part of this section.

Consider a binary series,  $s(t)$ , representing the occupancy in a specific area for  $t = 1, 2, \dots, N$ , where  $s(t) \in [0, 1]$ . This signal has two states of ‘on’ and ‘off’ representing the presence and absence from a specific area.

$$s(t) = (1, \dots, 1, 0, \dots, 0, 1, 1, \dots) \quad (1)$$

To have a more efficient form of presentation for the binary series, it is recommended to convert  $s(t)$  to a start-time and stop-time signal,  $x(t)$ , as stated below:

$$x(t) = (t_{s_1}, t_{e_1}, t_{s_2}, t_{e_2}, \dots, t_{s_i}, t_{e_i}, \dots, t_{s_n}, t_{e_n}) \quad (2)$$

where  $t_{s_i}$  and  $t_{e_i}$  are the start-time and stop-time of any event which has resulted in a value of 1 in  $s(t)$ . The new series  $x(t)$  in terms of the start/stop time has fewer values comparing with the long sequences in  $s(t)$ . This particular conversion provides an accurate and more flexible representation of binary series.

### 3.3. Start-time and Duration

Start-time and duration method is used when a binary signal is converted into two separate sequences of real numbers representing the start-time and duration of each activity respectively. It should be noted that there is a dependency between these two sequences.

Considering the binary series represented in expression 1, start-time series,  $y(t)$ , and duration series,  $d(t)$ , are represented respectively as:



$$y(t) = (t_{s_1}, t_{s_2}, \dots, t_{s_i}, \dots, t_{s_n}) \quad (3)$$

$$d(t) = (t_{e_1} - t_{s_1}, t_{e_2} - t_{s_2}, \dots, t_{e_n} - t_{s_n}) \quad (4)$$

This form of binary sequence conversion helps with the visualisation of the binary series extracted from occupancy sensors, since it is rather difficult to track the movements of the occupancy using large binary series. This form of representation will make sure that the time dependency and correlation within each component of the series is not lost.

The start-time and stop-time form of conversion has proved to be effective for modelling and prediction, while the start-time and duration form of conversion has proved to be more suitable for binary signal visualisation. The start-time and stop-time series,  $x(t)$  is normalised to a range between 0 and 1 or  $-1$  and 1 before it is applied to any network for modelling and prediction. The normalised signal  $x(t)$  will be used to represent the occupancy behaviour in an area. The next step is to establish a non-linear model representing this sequence.

#### 4. Prediction Model

In this paper, two types of recurrent neural network models are used to predict the data extracted from occupancy sensors. A short description of the models used, NARX and Elman, is presented.

##### 4.1. The NARX Network

Figure 3 shows the typical architecture of NARX network. In this model, Multi-Layer Perceptron (MLP) is used to approximate the following function [39].

$$y(t) = f(x(t-1), x(t-2) \dots x(t-D_x), y(t-1), y(t-2) \dots y(t-D_y)) \quad (5)$$

where  $x(t)$  and  $y(t)$  are respectively the input and the output of the model at time step  $t$ , while  $D_x$  and  $D_y$  are the input and the output memory orders with  $D_x \geq 1$ ,  $D_y \geq 1$  and  $D_y \geq D_x$ . Also,  $f$  is a non-linear function of the input and output of the model. The predicted output  $y(t)$  is regressed on the input value (exogenous)  $x(t-1)$  and the output value  $y(t-1)$  [40][41]. In this case, since one of the inputs of NARX is the output of the network,

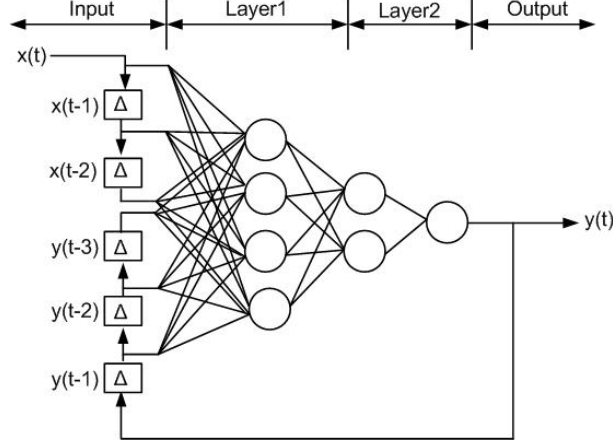


Figure 3: NARX Network with two input delays and three output delays.

this makes NARX network represent the dynamical characteristic of a system efficiently [42]. NARX network can be implemented as Time Delay Neural Network (TDNN) when its output memory order takes a zero value. Accordingly, there is only a time delay line in the input layer of the MLP learning algorithm which is used to approximate the following function in TDNN [43].

$$y(t) = f(x(t-1), x(t-2) \dots x(t-D_x)) \quad (6)$$

Generally, there are two modes in which NARX network is trained [42][44]. These modes are:

1. Parallel Mode: in this case, the estimated output of the network is returned to the input of the MLP. This mode is shown in Figure 4-a where TDL refers to the tapped delay line. This mode can provide a good estimation when one adds a regressive factor of the estimated value. As a result, the main dynamic features of the system are obtained.

$$y(t+1) = \hat{f}[(y(t), x(t))] \quad (7)$$

$$= \hat{f}[y(t-1), y(t-2), \dots y(t-D_y); \quad (8)$$

$$x(t-1), x(t-2), \dots x(t-D_x)]$$

2. Serial-Parallel Mode: in this case, instead of feeding back the estimated output of the network, the actual output is returned to the input of the neural network as illustrated in Figure 4-b. This mode has an important characteristic in that it has strictly feed forward architecture and a static back propagation learning algorithm can be used.

$$y(t+1) = \hat{f}[(\hat{y}(t), x(t))] \quad (9)$$

$$= \hat{f}[\hat{y}(t-1), \hat{y}(t-2), \dots, \hat{y}(t-D_y); \quad (10)$$

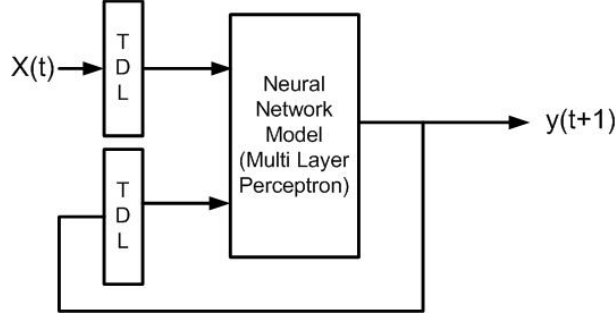
$$x(t-1), x(t-2), \dots, x(t-D_x)]$$

The embedded memory of NARX network gives a shorter path for gradient information in case the network is unfolded in time to back propagate the error signal. Having such a characteristic, the gradient descent learning is better in NARX networks at learning the long term dependencies. In gradient-based training algorithms, for  $n$  time steps in the past, the fraction of the gradient becomes zero as  $n$  increases. Vanishing gradient is a problem in other neural networks such as back propagation through time [41][40].

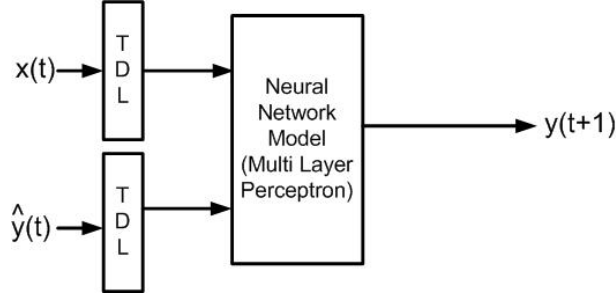
NARX network has been applied for modelling many real world applications, including time series and various artificial non-linear systems modelling [7]. In this paper, NARX network is applied to data collected from occupancy sensors to extract the behaviour pattern of an occupant. The Serial-Parallel architecture of NARX network is used for learning the dynamic behaviour of an occupant living in an intelligent environment. Then, the results obtained from NARX network are compared with simple recurrent neural network such as Elman network [45]. A brief description of the Elman recurrent neural network (also known as a partial recurrent network) is presented below.

#### 4.2. The Elman Network

Figure 5 shows the structure of the network with two inputs and one output unit. Elman network has three layers: input, hidden and output layers, with a number of “context units” in the input layer. Initially, the context units take zero values and then the output values of the hidden layer units at the previous time steps will be copied into these units. In this case, the network can perform time function mappings that are beyond the ability of the standard MLP. This characteristic allows Elman network to be suitable for time series prediction. Accordingly, the memory of the Elman network is constructed via a feedback. Using this feedback, temporal and spatial patterns can be learned, recognised and generated [46][47][48].



(a) Parallel network



(b) Serial-Parallel network

Figure 4: NARX network architecture.

## 5. Prediction Performance Measurement

To identify the ability of the prediction model, it is important to choose an appropriate error measure. In this study, depending on the input to the model, whether it is binary or non-binary, different error measurements are used.

For binary series, after training the binary time series data, a similarity measurement between the two binary series (i.e. actual and predicted data) is used. Different similarity coefficients have been proposed by researchers in different fields. A similarity coefficient indicates the degree of similarity between object pairs. In this investigation, these measures are used to find the similarity between the predicted and the actual datasets.

Jaccard coefficient [49] is one of the most commonly used similarity measurement coefficients for binary series. We have compared and investigated this measure in our study. Jaccard index can be used to find the mismatching bits between two binary vectors representing two different datasets. It is formally defined as follows:

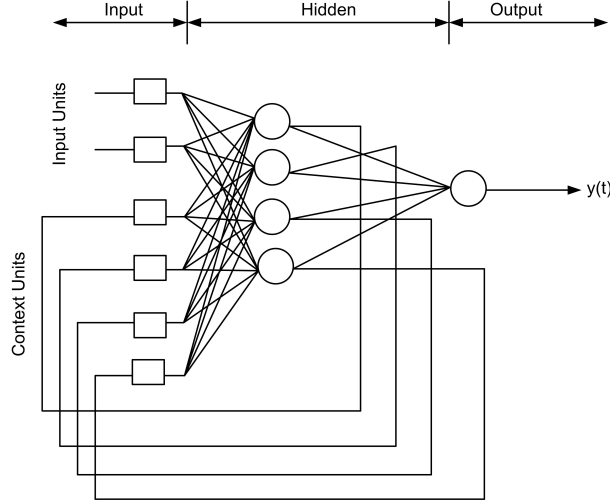


Figure 5: Elman network structure.

$$S_{ij} = \frac{a}{a + b + c} \quad 0 \leq S_{ij} \leq 1 \quad (11)$$

where  $S$  is the similarity between two binary set.  $a$  is the number of 1's in both datasets,  $b$  is the number of 0's in the first set and number of 1's in the second set and  $c$  is the number of 1's in the first set and the number of 0's in the second set.

The above error index will be used only for binary series prediction. However, when signals are converted into either start-time and stop-time or start-time and duration, then a continuous error measure is required. We have used Root Mean Square Error (RMSE) as defined below to measure the difference.

$$RMSE = \sqrt{\frac{1}{M} \sum_{n=1}^M (y(n+1) - \hat{y}(n+1))^2} \quad (12)$$

where  $y(n+1)$  is the actual value of the time series,  $\hat{y}(n+1)$  is the predicted value, and  $M$  is the number of steps that network model has to predict[50]. RMSE is used to compute the differences between the actual observation values and the predicted values which result from our model.

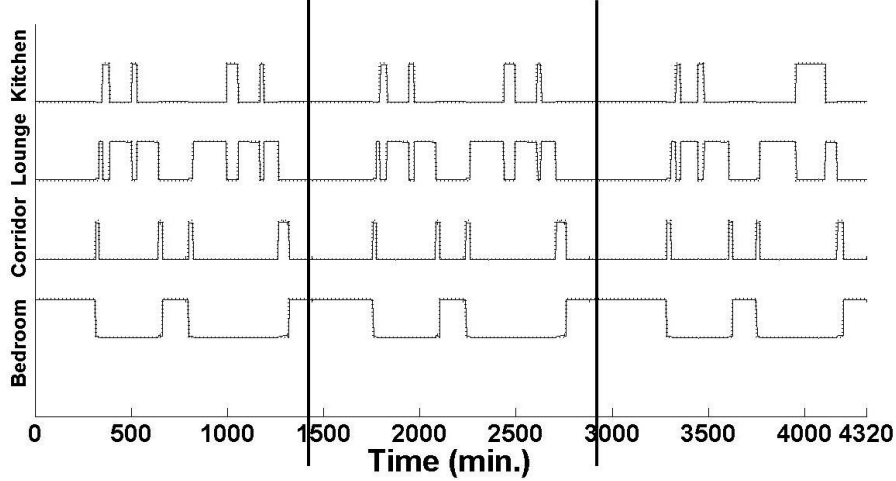


Figure 6: Two hours ahead predicted sensor values using NARX network for three days sample of simulated activities (-Predicted values; - - Actual values).

## 6. Experimental Results

To evaluate the performance of NARX network in modelling the binary series representing the occupancy pattern, various experimental works are conducted. In this section two separate datasets collected from a simulator, and a real environment are investigated. To evaluate the predictive ability of NARX network, the prediction results of NARX network is compared with Elman network.

A short summary of the data collection system is presented in the following section followed by the prediction results and the comparison.

### 6.1. Data Collection and Environment

The data collection system consists of an array of sensors, which collect information using Passive Infra-Red (PIR) motion sensors and door entry point contact switches. The occupancy/movement sensors and door switches are used to record the behaviour of the occupant. From these observations, the time that the occupant spent in one room before moving to another room is calculated.

Two separate datasets, simulated data generated from a simulator and real data collected from real environments, are investigated in our research.

- **Simulated Data:** We have already developed a simulator as reported in [51] to generate the occupancy signals (binary series). To generate a

simulated data, a number of uncertainty factors related to both modelled occupant as well as modelled environment are taken into accounts in modelling both the movement and duration. In this study, two different binary datasets are generated as if from occupancy sensors in the bedroom, corridor, lounge and kitchen with a resolution of one minute. The first dataset represents an occupancy signal for a very regular person with no duration uncertainty. The second dataset represents an occupancy signal for a less regular person with 6% duration uncertainty.

- **Real Data:** Binary data are collected from a real environment monitored by a JustChecking system [52]. Wireless sensors are installed in a house. These sensors are triggered as a person moves around the house. Data from the sensors are gathered by a controller and transferred to a web-server for further data analysis. Sensors are placed in the hallway, kitchen, lounge, the front/back door, etc. In this paper, front door, back door and four motion sensors including lounge, kitchen, bedroom and bathroom are used to detect movements of the person inside the house. The observations extracted from the occupancy sensors were recorded with a resolution of one second to monitor the daily behaviour of the occupant for a duration of over one year. It should be noted that collected data is based on a single occupant house. Logged data is time stamped and includes sensor ID and a sensor name (type). A sample of raw data collected from an intelligent environment is illustrated in Table 1. A software program is used to read the binary data from the actual occupancy sensors and transform them into a time series format containing the start/stop time and the status of the sensor.

## 6.2. Prediction Results

NARX and Elman networks are used to model and predict both simulated and real data as described earlier. The results are presented in two separate sections for simulated data and real data. To train these networks, experiments are conducted using a number of parameters which are necessary for training. These parameters are: number of input tapped delay  $D_x$ , number of feedback output delay  $D_y$ , and number of neurons in the hidden layer  $N$ .

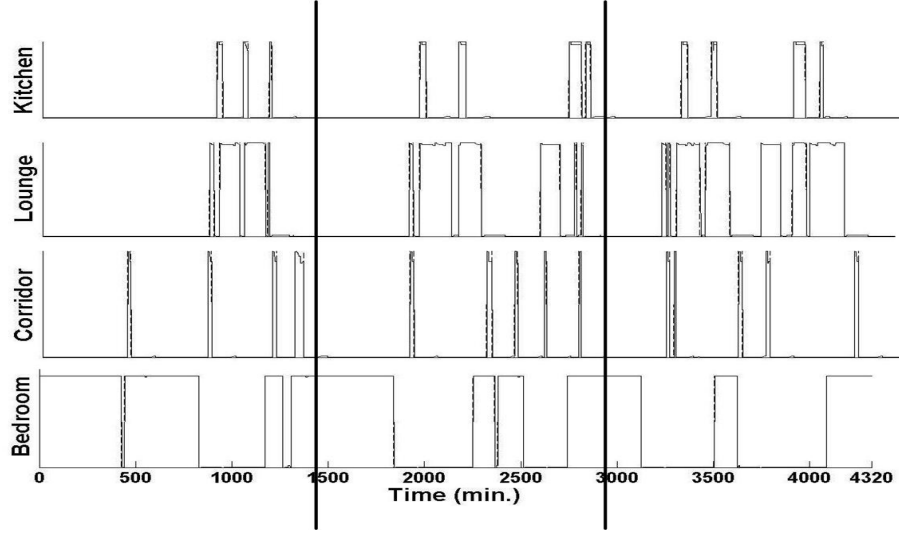


Figure 7: Two hours ahead predicted sensor values using NARX network for three days sample of simulated activities for an occupant with less regular activities (6% duration uncertainty) (- Predicted values; - - Actual values).

Table 1: A sample of collected raw data from the environment

Time Stamp	Sensor ID	Type
21/02/2007 01:15	5	Bedroom
21/02/2007 01:18	5	Bedroom
21/02/2007 01:18	7	Lounge
21/02/2007 01:19	7	Lounge
21/02/2007 01:19	8	Kitchen
21/02/2007 01:19	1	Front Door - open
21/02/2007 01:20	2	Front Door - close
21/02/2007 01:20	8	Kitchen
21/02/2007 01:21	8	Kitchen



### 6.2.1. Simulated Data Results

Ten days of binary data are used to evaluate the predictive ability of NARX network where the sampling interval is one minute. The length of the data is 14400 samples; we chose the first 10080 points for training the network and the last 4320 points for testing. Generally, for any ANN the prediction results depend on the kind of inputs of the network, whether they have binary or real values. Therefore, in this study, two sets of input data are tested; these are binary and real time series data. For binary input data, the input to the NARX network is driven by the actual binary time series extracted from the sensor data at time  $t$ . The output is the future value of the data, i.e. at the time  $t + \tau$  ( $\tau$  is the number of step ahead prediction).

To show the ability of NARX network in prediction of binary time series generated from the simulated data, we have conducted two sets of experiments on the simulated data. The first set of simulated data represent the binary time series for a very regular occupant i.e. the generated simulated signal has no duration uncertainty. The second set of data represent the binary time series for an occupant with less regular behaviour i.e. the generated signal has 6% duration uncertainty.

Figure 6 and Figure 7 show the NARX network prediction results for a sample of three days of the binary time series data with regular and irregular patterns respectively. For both of these figures, the solid line is the predicted values and the dotted line is the actual values. In addition, the prediction is based on two hours step ahead. In Figure 6, the predicted data is very close to the actual data in terms of durations since these data represent a very regular occupancy behaviour. For instance, the bedroom occupancy signal shows the sleeping patterns for an occupant who goes to his/her bed at a regular time (around 22 : 00) to sleep about 6 – 7 hours and also takes a nap of about 2 hours after mid day. The results of occupancy prediction using NARX network for two hours step ahead of the simulated data with some irregular patterns are shown in Figure 7. From this figure it can be observed that NARX network can successfully predict the irregular occupancy patterns where the occupant spent more time in some area than others in the environment. For example, the bedroom occupancy signal shows that the occupant spent most of his/her time in the bedroom which might indicate an early detection of illness or depression. However, the difference between the predicted and the actual data is barely noticeable. As stated in Section 5, the errors between the predicted and the actual data for the binary data is calculated using some binary similarity measurements.

In this work, Jaccard coefficient is applied to calculate the similarities

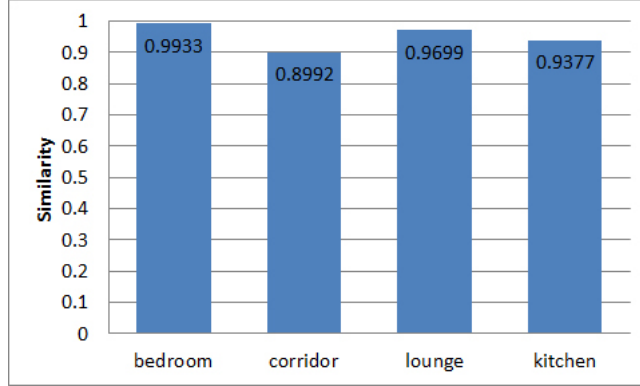
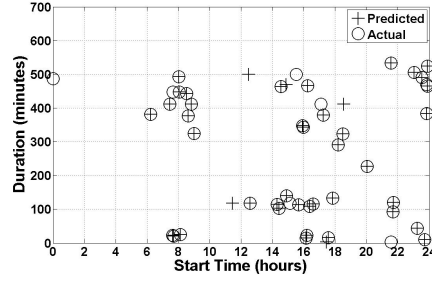


Figure 8: Comparing similarity measures between the actual and the predicted binary sensor values of the results shown in Figure 7.

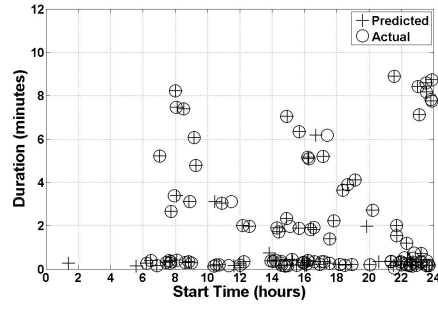
between the predicted and actual datasets. For the simulated data with some regular patterns, the similarity between the predicted and the actual binary datasets for the four motion sensors (bedroom, corridor, lounge and kitchen) are ranged between 99% and 98%. Similarity measure using the Jaccard coefficient were almost close to 1, indicating that the differences between the actual and the predicted output of the NARX network are negligible. For the simulated data with irregular patterns, Figure 8 shows the similarities between these two binary datasets. The similarity between the predicted and the actual binary datasets for the four motion sensors are ranged between 99% and 89%. We think that Jaccard coefficient is a good measure of prediction error. It excludes the negative matches considering only positive matches and mismatched bits between two binary vectors. For example, in Figure 6, for the most time the corridor motion sensor binary dataset has a zero value and only on a few occasions does it have a value equal to one.

Therefore, it is apparent that even using data with some irregular patterns, we would still be able to predict the movement activities of an occupant living in an intelligent environment.

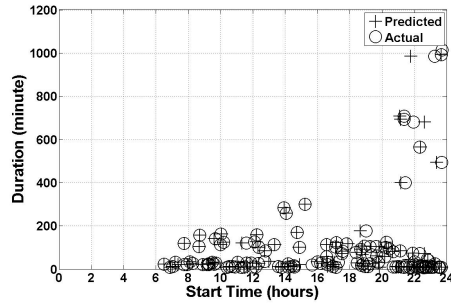
Additionally, we conducted experiments where the inputs to the NARX network were the sensor data after conversion from binary to start-time and stop-time method. In this case, the inputs to the NARX network are the converted data using the start-time and the stop-time at time  $t$ , while the output is the start-time and the stop-time at the time  $t + \tau$ . Figure 9 shows the predicted results of the second set of the simulated data with 6% uncertainty using NARX network to predict one step ahead using different  $D_x$  and



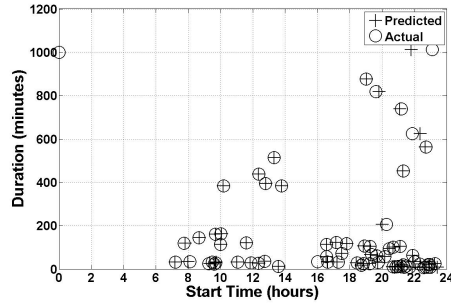
(a) Bedroom motion sensor



(b) Corridor motion sensor



(c) Lounge motion sensor



(d) Kitchen motion sensor

Figure 9: The prediction results for four sensors representing a simulated environment with 6% uncertainty based on start/duration time.

Table 2: The prediction results of the lounge sensor for Elman and NARX network using different number of feedback outputs (D).

No. of Hidden units	Elman Network	NARX Network
D=1	3.4753	0.0854
D=3	4.0820	0.0517
D=10	0.2027	0.0418

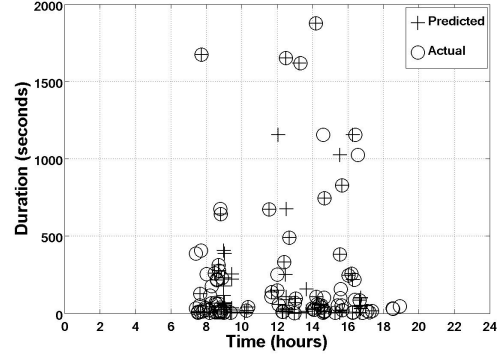
$D_y$  values for each time series data. In these graphs, start-time/duration method is used to display the actual and predicted values. Start-time and duration method of conversion seems to be a practical and easy tool to visualise large binary data. To measure the accuracy of prediction based on start-time and stop-time method of conversion, RMSE is used. The RMSE for both training and testing of all simulated sensor data range from minimum of 5% to maximum of 9%. In Figure 9 the  $\circ$  are the predicted values using NARX model and the  $+$  are the actual sensor values.

### 6.2.2. Real Data Results

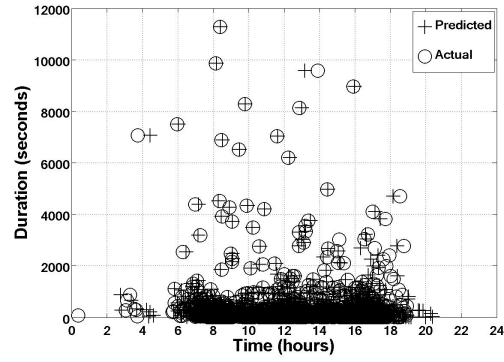
The NARX model is also exploited to test real data in order to monitor the movement of a person in his/her home. The real data are split into a sample of 14 days for training (i.e. about 1,209,600 points of binary datasets) and 6 days for testing (i.e. about 518,400 points of binary time series). These datasets are trained using NARX network with a different number of input and output memory orders and hidden units. Considering the size of the real dataset we were not able to generate a model based on the actual binary data. Therefore, we only use the start-time and stop-time converted dataset. RMSE to train NARX network for all real datasets are ranging from 6 – 9%. The predicted results of the back-door sensor, the lounge and the kitchen motion sensor are illustrated in Figure 10. In these experiments, one step ahead (next start/stop time cycle) is predicted. From Figure 10-a, it is obvious that the back door is usually opened about 7 minutes or a bit more, and it is sometimes left open for more than 20 minutes. In Figure 10 the  $\circ$  are the predicted values using NARX model and the  $+$  are the actual sensor values.

### 6.3. Comparison with Benchmark Models

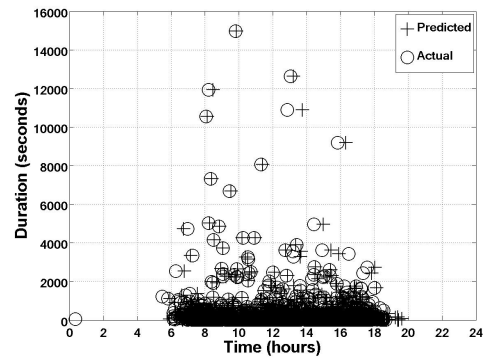
To show the validity and significance of the results, our prediction results are compared against the results from some benchmark models.



(a) Back door sensor



(b) Lounge motion sensor



(c) Kitchen motion sensor

Figure 10: The prediction results for three sensors representing a real environment.

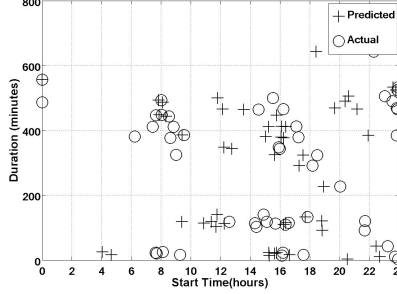


Figure 11: The prediction results for the bedroom sensor representing a simulated environment with 6% uncertainty based on start/duration time using moving average model.

### 6.3.1. Comparison with Moving Average

A simple naive approach uses the previous value of a time series to predict the next value. Three averaging forecasting techniques are available. They are: moving average, weighted moving average and exponential smoothing. Moving average approach is used and compared with the predicted results of the NARX network stated in the previous section. For instance, the RMSE for the first set of the simulated data (i.e. for a very regular occupant) are 4.4746, 4.4489, 0.2562 and 2.7765 for the bedroom, corridor, lounge and kitchen sensors data respectively. The RMSE of all simulated sensor data using NARX network was ranging between 5 – 9% for the four sensors data sets as stated in Section 6.2.1. The prediction is based on the start/stop time as inputs to the networks to predict the next start/stop time. The prediction results of NARX network are much better than moving average approach since it presents the lowest prediction errors.

Further experiments are also conducted on the second set of the simulated data; again the moving average results performed poorly. The predicted values for bedroom sensor data using moving average is shown in Figure 11. Comparing this figure with Figure 9-a where NARX network is used show a much better prediction results. There are clear differences between the predicted and actual sensor values appear in Figure 11 when moving average approach is used. Our observation is that the prediction of human behavioural patterns using naive approaches is hard since these approaches work well when the data is quite stable over time.

### 6.3.2. Comparison with Elman Network

The results of the NARX network presented above are compared with another RNN namely Elman network. Both networks are trained using

Table 3: Prediction error for the lounge motion sensor based on a sample of ten days training data with different values for step ahead prediction (step represents the next start/stop time cycle of the occupant movement activity).

Number of step ahead	RMSE
1 step*	0.0765
2 step	0.0697
3 step	0.0977
4 step	0.5783
5 step	0.1134

the MLP algorithm with the same number of hidden units. Once again, experiments are conducted on the simulated data and the real data. For example, the prediction results for the simulated lounge sensor data are summarised in Table 2. The prediction is based on the start/stop time as inputs to the networks. In this experiment, the number of the hidden units play an important role in network convergence. Increasing the number of the hidden units results in decreasing the RMSE of training for both networks. However, the training time and number of epochs that are required to train the network increases rapidly in Elman network in contrast to NARX. For instance, training the lounge sensor data using Elman network requires 1000 epochs for training, while only 50 epochs are required using NARX network.

The NARX network is also compared with the Elman network using the real data. Based on our experiments, again the convergence of NARX network is much better than the Elman network to train the real data. For example, the RMSE for training the back-door and front-door sensor data using Elman network are 1.75 and 2.15 respectively. However, the RMSE is only 0.05 and 0.07 to train the back-door and front-door using NARX network.

#### 6.4. Discussion

This study set out with the aim of assessing the application of a recurrent NARX network in identification and prediction the behaviour patterns of an occupant living in an intelligent environment. In the current study, the NARX network is applied for both the binary sequences and the converted data into start/stop time sequences. As explained in Section 3, the start/stop time approach compresses the long binary series. Although the NARX network gives good results in predicting the binary inputs, the complexity of the structure of NARX network to train the converted data is

far lower than when training the original, unconverted binary time series. Fewer hidden neurons and a small number of epochs of training are used to train the network.

Experimental results also showed that as the number of hidden neurons are increased the convergence becomes better although the training time is increased. For instance, consider training the data collected from the kitchen motion sensor of the real environment where the number of input and output memory orders are 100. In this case, the RMSE are 0.0319 for 20 hidden neurons and 0.0263 for 50 hidden neurons. Increasing the number of output memory order is a major factor leading to much better convergence. From our experiments on the sensor data using NARX and Elman networks, it can be concluded that the NARX network performs better than the Elman network.

The capability of the system to predict the occupancy behaviour was assessed by investigating the effect of increasing the number of step ahead  $\tau$  in network performance convergence. Therefore, different experiments are carried out on binary series. The results of this study show that, as  $\tau$  is increased for a number of steps, the RMSE is also increased. The increasing amount depends on the type of the sensor, number of hidden neurons and the number of input/output memory order. Table 3 summarises the NARX network prediction results for the lounge motion sensor dataset generated from the simulated data with 6% duration uncertainty for a period of ten days. In this experiment, two hidden layers with ten neurons in each layer are used. Also, two input/output memory orders are used. From the results in Table 3 it can be concluded that the increase in prediction time decreases the accuracy of the network convergence. More specifically, the RMSE is increased after just 3 steps ahead prediction. It should be noted that in this table, step means the next start-time and stop-time cycle of the occupant movement activity.

The experimental results were successful as the network was able to recognise the behavioural patterns of an occupant that might be indicators of developing physical or mental medical conditions. Also, the results of this study indicates that NARX network gives good results in predicting the binary time series data collected from an intelligent environment.

## 7. Conclusions

This paper shows that the occupancy pattern extraction and prediction in an inhabitant intelligent environment can be efficiently modelled using recurrent neural networks such as NARX network. This network has been



used to predict the binary data sequence extracted from occupancy sensors in the simulated and real environments. The predicted results of NARX network are tested and compared with other recurrent neural networks such as Elman network. Due to the feedback from the output layer to the input layer with a proper number of input time delays, NARX network gives better results than Elman network. In addition, NARX network has shown its ability in a long-term prediction compared with other neural network predictive algorithms.

In addition, this paper highlights the need for a flexible and efficient data representation and visualisation in large binary datasets. The start-time and stop-time approach is used to convert and represent the binary datasets. The result of this conversion is a compressed data compared with the original long series of binary data items, which are then used as inputs to our prediction techniques. The complexity of the structure of NARX network to train the data after conversion from binary to start-time and stop-time is lower than when training the original, unconverted binary time series. Fewer hidden neurons and a small number of epochs of training are used to train the network.

Further work in this area needs to explore the use of other techniques such as the kernel learning methods suggested in [27] to improve the performance of handling large sparse data sets. Additionally, extension to the length of the predicted pattern is essential to make the developed system suitable for monitoring the elderly in these circumstances.

## References

- [1] G. Virone, A. Sixsmith, Activity prediction for in-borne activity monitoring, IET 4th International Conference on Intelligent Environments (2008) 1–4.
- [2] C. Doukas, I. Maglogiannis, Advanced patient or elder fall detection based on movement and sound data, in: Proc. of the Second International Conference on Pervasive Computing Technologies for Healthcare, pp. 103–107.
- [3] D. N. Monekosso, P. Remagnino, Behavior analysis for assisted living, IEEE Transactions on Automation Science and Engineering 7 (4) (2010) 879–886.
- [4] D. Cook, S. Das, Smart Environments: Technology, Protocols and Applications, Wiley, 2004.
- [5] L. Atallah, B. Lo, Y. Guang-Zhong, F. Siegemund, Wirelessly accessible sensor populations (wasp) for elderly care monitoring, in: Proc. of the Second International Conference on Pervasive Computing Technologies for Healthcare, 2008, pp. 2–7.
- [6] A. Lotfi, C. Langensiepen, S. M. Mahmoud, M. J. Akhlaghinia, Smart homes for the elderly dementia sufferers: Identification and prediction of abnormal behaviour, Journal of Ambient Intelligence and Humanized Computing 3 (3) (2012) 205–218.
- [7] L. Medsker, L. Jain, Recurrent neural networks: design and applications, The CRC Press International Series on Computational Intelligence.
- [8] L. Atallah, G.-Z. Yang, The use of pervasive sensing for behaviour profiling - a survey, Pervasive and Mobile Computing 5 (5) (2009) 447–464.
- [9] D. N. Monekosso, P. Remagnino, Anomalous behaviour detection: Supporting independent living, Advanced Information and Knowledge Processing, Springer (2009) 33–48.
- [10] T. V. Duong, H. H. Bui, D. Q. Phung, S. Venkatesh, Activity recognition and abnormality detection with the switching hidden semi-markov model, in: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, 2005, pp. 838–845.

- [11] F. Cardinaux, S. Brownsell, M. Hawley, Modeling of behavioural patterns for abnormality detection in the context of lifestyle reassurance, 2008, pp. 243–251.
- [12] H. Li, Q. Zhang, P. Duan, A novel one-pass neural network approach for activities recognition in intelligent environments, in: Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on, 2008, pp. 50–54.
- [13] S.-H. Park, J.-H. Lee, J.-W. Song, T.-S. Park, Forecasting change directions for financial time series using hidden markov model, Lecture Notes in Computer Science, Springer 5589 (2009) 184–191.
- [14] O. Yogurtcu, E. Erzin, A. Gursoy, Extracting gene regulation information from micro array time-series data using hidden markov models, Lecture Notes in Computer Science, Springer 4263 (2006) 144–153.
- [15] C.-C. Tseng, D. Cook, Mining from time series human movement data, IEEE International Conference on Systems, Man and Cybernetics 13 (2006) 3241–3243.
- [16] L. Vintan, A. Gellert, J. Petzold, T. Ungerer, Person movement prediction using neural networks, in: In First Workshop on Modeling and Retrieval of Context, 2004.
- [17] D. J. Cook, Making sense of sensor data, IEEE Pervasive Computing (6) (2007) 105–108.
- [18] M. Mozer, The neural network house: An environment that adapts to its inhabitants, 1998, pp. 110–114.
- [19] H. Zheng, H. Wang, N. Black, Human activity detection in smart home environment with self-adaptive neural networks, IEEE International Conference on Networking, Sensing and Control (2008) 1505–1510.
- [20] R. J. Frank, N. Davey, S. P. Hunt, Time series prediction and neural networks, J. Intell. Robot. Syst. 31 (2001) 91–103.
- [21] G. P. Zhang, V. L. Berardi, Time series forecasting with neural network ensembles: an application for exchange rate prediction, Journal of the Operational Research Society 52 (2001) 652–664.
- [22] M. Han, J. Xi, S. Xu, F.-L. Yin, Prediction of chaotic time series based on the recurrent predictor neural network, Signal Processing, IEEE Transactions on 52 (12) (2004) 3409–3416.

- [23] A. Hussain, S. Squartini, F. Piazza, Recurrent multi-scale architecture for long-term memory prediction task, *Speech, and Signal Processing* 2 (6-10) (2003) 789–792.
- [24] R. Illingworth, V. Callagha, H. Hagra, Towards the detection of temporal behavioural patterns in intelligent environments, 2006, pp. 119–125.
- [25] R. Illingworth, V. Callagha, H. Hagra, A neural network agent based approach to activity detection in ami environments, 2005, pp. 1–12.
- [26] Y. Liu, W. L. Chen, Z. L. Gao, H. Q. Wang, P. Li., Adaptive control of nonlinear time-varying processes using selective recursive kernel learning method, *Industrial and Engineering Chemistry Research* 50 (2011) 2773–2780.
- [27] Y. Liu, H. Wang, J. Yu, P. Li, Selective recursive kernel learning for online identification of nonlinear systems with narx form, *Journal of Process Control* 20 (2) (2010) 181–194.
- [28] K. Vrotsou, K. Ellegard, M. Cooper, Everyday life discoveries: Mining and visualizing activity patterns in social science diary data, *Information Visualization*, 11th International Conference (2007) 130–138.
- [29] R. Kadouche, H. Pigot, B. Abdulrazak, S. Giroux, User’s behaviour classification model for smart houses occupant prediction, *Atlantis Ambient and Pervasive Intelligence* 4 (2011) 149–164.
- [30] S. K. Das, D. J. Cook, A. Bhattacharya, E. O. Heierman, T. Y. Lin, The role of prediction algorithms in the mavhome smart home architecture, *IEEE Wireless Communications* 9 (2002) 77–84.
- [31] B. S. Lee, T. P. Martin, N. P. Clarke, B. Majeed, D. Nauck, Dynamic daily-living patterns and association analyses in tele-care systems, in: *Proc. of the Fourth IEEE International Conference on Data Mining*, Nov. 2004, pp. 447–450.
- [32] T. Martin, B. Majeed, B.-S. Lee, N. Clark, Fuzzy ambient intelligence for next generation telecase, in: *Proc. of the IEEE Int. Conf. on Fuzzy Systems*, Vancouver, Canada, 2006, pp. 894–901.
- [33] A.-M. Vainio, M. Valtonen, J. Vanhala, Proactive fuzzy control and adaptation methods for smart homes, *IEEE Intelligent Systems* 23 (2) (2008) 42–49.

- [34] H. Hagaras, V. Callaghan, M. Colley, G. Clarke, A. Pounds-Cornish, H. Duman, Creating an ambient-intelligence environment using embedded agents, *IEEE Intelligent Systems* 9 (6) (2004) 12–20.
- [35] D. Furno, V. Loia, M. Veniero, M. Anisetti, V. Bellandi, P. Ceravolo, E. Damiani, Towards an agent-based architecture for managing uncertainty in situation awareness, in: *Proc. of the 2011 IEEE Symposium on Intelligent Agent*, April 2011.
- [36] H. Medjahed, D. Istrate, J. Boudy, B. Dorizzi, Human activities of daily living recognition using fuzzy logic for elderly home monitoring, *FUZZ IEEE Korea* (2009) 20–24.
- [37] M. J. Akhlaghinia, A. Lotfi, C. Langensiepen, N. Sherkat, Occupant behaviour prediction in ambient intelligence computing environment, *International Journal of Uncertain Systems* 2 (2) (2008) 85–100.
- [38] F. Doctor, H. Hagaras, V. Callaghan, A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments, *IEEE Transaction on Systems, Man, and Cybernetics-part A: Systems and Humans* 35 (2005) 55–65.
- [39] J. Sum, W.-K. Kan, G. Young, A note on the equivalence of narx and rnn, *Neural Computing and Applications* 8 (1) (1999) 33–39.
- [40] J. M. P. Menezes, G. A. Barreto, Long-term time series prediction with the narx network: An empirical evaluation, *Elsevier Science* 71 (16-18) (2007) 3335–3343.
- [41] M. Siek, D. Solomatine, Chaotic model with data assimilation using narx network, 2009, pp. 2516–2523.
- [42] C. Jiang, F. Song, Forecasting chaotic time series of exchange rate based on non-linear autoregressive model 5 (2010) 238–241.
- [43] T.-N. Lin, C. L. Giles, B. G. Horne, S.-Y. Kung, A delay damage model selection algorithm for narx neural networks, *Signal Processing, IEEE Transactions on* 45 (11) (1997) 2719–2730.
- [44] H. Xie, H. Tang, Y.-H. Liao, Time series prediction based on narx neural networks: An advanced approach, in: *Proc. of the International Conference on Machine Learning and Cybernetics*, Vol. 3, 2009, pp. 1275–1279.

- [45] J. L. Elman, Finding structure in time, *Cognitive Science* 14 (2) (1990) 179–211.
- [46] A. I. Galarniotis, A. C. Tsakoumis, P. Fessas, S. S. Vladov, V. M. Mladenov, Using elman and fir neural networks for short term electric load forecasting, in: *Proc. of the International Symposium on Signals, Circuits and Systems*, Vol. 2, 2003, pp. 433–436.
- [47] S. Vitabile, M. Pernice, S. Gaglio, Daily peak temperature forecasting with elman neural networks, in: *Proc. of the IEEE International Joint Conference on Neural Networks*, Vol. 4, 2004, pp. 2765–2769.
- [48] W. Y. Goh, C. P. Lim, K. K. Peh, Predicting drug dissolution profiles with an ensemble of boosted neural networks: a time series approach, *IEEE Transactions on Neural Networks* 14 (2) (2003) 459–463.
- [49] Y. Yin, K. Yasudab, Similarity coefficient methods applied to the cell formation problem: A taxonomy and review, *Int. J. Production Economics* 101 (2006) 329–352.
- [50] A. Chitra, S. Uma, An ensemble model of multiple classifiers for time series prediction, *International Journal of Computer Theory and Engineering* 2 (3) (2010) 1793–8201.
- [51] M. J. Akhlaghinia, A. Lotfi, C. Langensiepen, N. Sherkat, Occupancy simulator for a single-occupant ambient intelligent environment, in: *Proc. of the 7th IEEE International Conference on Cybernetic Intelligent Systems*, 2008, pp. 939–946.
- [52] JustChecking, Supporting independence people with dementia, ([www.justchecking.co.uk](http://www.justchecking.co.uk)) (January 2012).