

Consumer Load Prediction Based on NARX for Electricity Theft Detection

Abdullateef Ayodele Isqeel

Department Electrical/Electronic Engineering
Yaba College of Technology
Lagos, Nigeria
e-mail: Isqeel@hotmail.co.uk

Salami Momoh-Jimoh Eyiomika

Department of Mechatronic Engineering
International Islamic University
Gombak, Malaysia
e-mail: momoh@iium.edu.my

Tijani Bayo Ismaeel

Department of Electronic/Electrical
ADMC, HCT
Abidabi, UAE
e-mail: tijani.ismail@hct.ac.ae

Abstract— A range of load prediction techniques has largely been used for energy management at various levels. However, the data used for the prediction are cumulative energy data, which reveal the activities of consumers and not individual consumer, on the distribution power network. Individual consumer data is essential for real time prediction, monitoring and detect of electricity theft. A new approach of monitoring individual consumer based on consumer load prediction using nonlinear autoregressive with eXogenous input (NARX) network is considered in this study. One month average energy consumption data acquired from consumer load prototype developed was used. Consequently, 5-minute step ahead load prediction was achieved. The NARX architecture was based on nine hidden neurons and two tapped delay and the network trained using Bayesian regulation backpropagation technique. The data set contains a total of 8928 data points representing energy consumed at five minute interval for one month. The data was divided into two sets at ratio 70:30 for training and validation, respectively. The training data equals 6206 while the validation data is 2722. MATLAB environment was used for the processing of the data. The training and validation MSE is 0.0225 and 0.0533 respectively, while the total time for the training is 0.016s.

Keywords- Nonlinear autoregressive with eXogenous input, consumer load prediction, energy management, data acquisition, Electricity theft detection, Power distribution network, Bayesian backpropagation technique.

I. INTRODUCTION

Consumer load prediction is important to power system management. Besides, facilitating the estimation of future electricity demand by the utilities based on the past records, it also helps the utilities in decision making and operational planning such as routine maintenance, dispatch of generators, management of fuel and development of infrastructure[1]. Inaccurate load prediction increases the operation cost of the utility company most especially under a deregulated market environment [2]. In the last decade, very short term load forecasting (VSTLF) which predicts load consumption from few minute to several hours has been proposed [3]. VSTLF is used in energy management

for load frequency control, dispatch functions and areas that need forecast for short time leads such as real time consumer load prediction which can be used for monitoring individual consumer. Historical load data, weather data and seasonal data are the common input variables in load forecasting. However, instead of modeling the relationships between these variables, as in other types of load prediction, VSTLF extrapolates the recently observed load pattern to the nearest future [4] as its relied on real time data to predict future energy demand.

Generally, aggregate consumers' data have been used for electric power load prediction [5]. These data are short of revealing individual consumer energy pattern, rather it reveals cumulative energy pattern which indicates activities of the consumers connected to the network. Individual consumer data is essential for load prediction as well as monitoring of individual consumer's activities such as electricity theft on the distribution network. The consequences of electricity theft has been well reported in terms of enormous loss of revenue, properties and lives on the part of the energy providers and consumers [6, 7]. In order to reduce the rate of stealing of electricity, the load activities of individual consumer has to be monitored and predicted. Thus, NARX network is proposed for the consumer load prediction which will be bench marked with the actual load to detect electricity theft on the distribution network.

II. LITERATURE REVIEW

Various techniques of predicting electricity load have been reported, however the techniques on very short term load prediction are limited. One of these techniques is statistical method which include multiple linear regression, exponential smoothing, time series, state space, and Kalman filter techniques [8]. Taylor [9] evaluated various methods of very short term load prediction between 10 and 30 minutes ahead using minute-by-minute British electricity demand load data.. Singh, et al. [10] proposed seasonal autoregressive moving average (SARMA) for home peak load and the study noted that ability to predict the stochastic activities of the consumer and routines are more significant for home load prediction. Weekday based prediction model for electricity load forecasting using autocorrelation feature selection and machine learning algorithm, proposed by Koprinska, et al. [11], are

considered as global model that predicts the load for all days of the week and local model that predict load for each day of the week. The two models were used to analyze a 5-minute electricity data obtained within two years from the State of New South Wales in Australia and the performances of the models were relatively good. Trudnowski, et al. [12] proposed a strategy for developing a very short-term load prediction using slow and fast Kalman estimators and an hourly forecaster load prediction for power system automatic generation control. The Kalman model parameters were determined by matching the frequency response of the estimator to load residuals.

The other techniques to VSTLF is the application of various artificial intelligent paradigm such as artificial neural network (ANN), fuzzy logic (FL) and knowledge based expert systems [13-16]. A method of very short-term loads forecasting using wavelet neural networks with data pre-filtering 1-h into the future in 5-min steps was proposed [17]. Yang, et al. [18] proposed fuzzy neural system (FNS) for very-short-term electric load prediction based on chaotic dynamics reconstruction technique. Hybrid model of similar day and neural network for load forecast ranging from 15 minutes to few hours was proposed by Fok and Vai [19]. Hourly weather information which is not often considered in other VSTLF literatures was considered as one of input variables, while the results of VSTLF was used to adjust a day ahead STLF.

The detailed review of several techniques proposed and developed for the detection and estimation of electricity theft have been reported by Abdullateef, et al. [6]. These techniques based on the energy consumption pattern of the consumer include, automatic meter reading system [20-22] and Support vector machine (SVM), [23] [7]. Others include power line impedance technique proposed by Pasdar and Mirzakuchaki [24], central observer meter (COM) to monitor and identify the perpetrators was proposed by Bandim, et al. [25]. The method proposed by Cavdar [26], uses two energy meters to track illegal connection. Meter tampering detection based on changes between live and neutral current as well as voltage monitoring at the meter input terminals was proposed by Naiman, et al. [27]. The injection of unwanted harmonics into the distribution network in order to cause damage to the appliances of the suspected illegal users and application of smart resistance incorporated in smart meter as a mode of detecting illegal electricity usage, have been reported by Bat-Erdene, et al. [28] and Kadurek, et al. [29] respectively.

III. METHODOLOGY

A. NARX network

Nonlinear autoregressive network with eXogenous inputs (NARX) is a recurrent dynamic network usually used for input-output modelling of nonlinear systems such as time series [16]. There are many applications of NARX neural networks though, however, its application in system identification and time series prediction are most popular [30-32]. Aside the fact that NARX neural network can handle nonlinear data, its learning ability with gradient descent algorithms have been shown to be effective and

superior than any other recurrent network [33]. Furthermore, NARX network generalizes better and converges faster than other networks [34].

B. NARX Network Architecture

NARX network is formed by two tapped-delay lines, the input tapped delay which slides over the input signal and the output tapped delay which slides over the output signal. Its feedback network comes from output neuron only. This distinguishes it from other recurrent neural network that has feedback from hidden neurons. A typical NARX network is shown in Figure 1. Furthermore, it could be designed as a feed-forward time delay neural network (Without the feedback loop of delayed outputs) when applied to time series prediction.

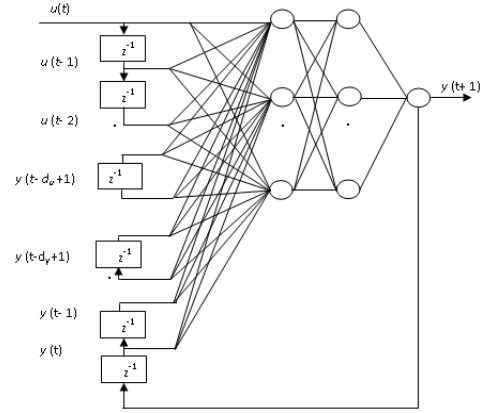


Figure 1 NARX Architecture with Input d_u and Output d_y Tapped Delays

NARX model can be expressed mathematically as

$$y(t+1) = f[y(t), y(t-1), \dots, y(t-d_y+1); u(t), u(t-1), \dots, u(t-d_u+1)] \quad (1.1)$$

Concisely it can be written as

$$y(t+1) = f[y(t); u(t)] \quad (1.2)$$

where $u(t) \in \mathcal{R}$ and $y(t) \in \mathcal{R}$ represent the input and output of the model at time t , $d_u \geq 1$ and $d_y \geq 1$, $d_u \leq d_y$, are the input memory and output memory orders respectively, which is analogous to the order of an autoregressive model. The function $f[\cdot]$ is a nonlinear mapping function which is approximated by using multi-layer perceptron neural network (MLP). The computational capabilities of NARX recurrent network as compared to other recurrent neural networks have been evaluated [35] and likened to a tuning machine.

However, NARX model performance depends upon the dimension of embedded memory of input and output and the number of neurons in the input layer. Lack of accurate method of choosing and optimizing the embedded memories in an optimal way rather than experimental (trial and error) method is one of the drawbacks of the NARX network [16]. The effects of inappropriate memory order in NARX network could hamper its learning ability as well as generalization performance [36]. One of the techniques in selecting appropriate embedded memory proposed by Tsung-Nan, Giles, Horne, and Sun-Yuan (1997) attempts to select optimal memory orders of NARX networks by calculating the sensitivity of the error to each memory order after the network is trained by gradient-based learning algorithm.

C. NARX Training Modes

There are two types of training mode employed by NARX network; the parallel mode and the series-parallel mode. In parallel mode the estimated output is fed back into the input of the feedforward neural network as part of the NARX architecture. On the other hand, in the series-parallel mode, there is no feedback from the output, rather, the true output is used during the training instead of the estimated output. The advantages of the series-parallel is that the resulting network has a purely feedforward architecture, and static backpropagation can be used for training. Furthermore, the input to the feedforward network is more accurate due to the absence of estimated value.

IV. SELECTION OF NARX ARCHITECTURE

One of the challenges of using NARX network in prediction is selection of appropriate network parameters most especially the number of the hidden neuron and the tapped delay. These are essential for the complexity or otherwise of the network and are conventionally done experimentally (trial and error). Moreover, the training algorithm equally plays an important role in the architecture of the NARX. However, since the most popularly acceptable training algorithm is Bayesian regulation backpropagation, it was used for the training of the network. The hidden neuron was choosing by evaluating various number of neuron from one to ten and choosing the minimum based on MSE. The number of hidden neurons corresponding to the minimum MSE is nine and it was used for the consumer load prediction see Figure 2. In addition the two tapped delay was selected by trial and error for the prediction.

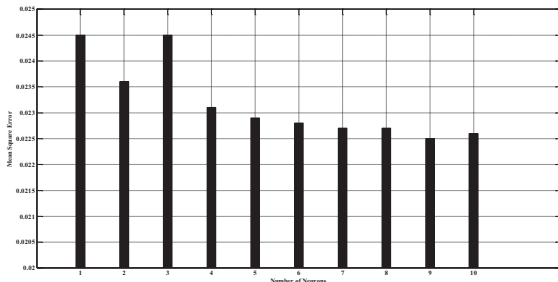


Figure 2: Conventional NARX Network Hidden Neuron Selection

V. DATA

The data used for the load prediction was acquired from the CLP [37], continuously for one month and it represents the consumer energy consumption over this period. The data set contains a total of 8928 data points representing energy consumed at five minute interval and it was divided into two sets at ratio 70:30 for training and validation, respectively. The training data equals 6206 while the validation data is 2722. MATLAB environment was used for the processing of the data and the results for a typical week are shown in Figure 5.7 to Figure 5.13. For clarity sake the x-axis time is scaled continuously in

hours throughout the month and span form 1 hour to 517 hour for the training and 1 hour to 227 hour validation. Also, the results presented one week from Sunday to Saturday with both training and validation results corresponding to one day data. The training and validation MSE is 0.0225 and 0.0533 respectively, while the total time for the training is 0.016s.

VI. RESULTS AND DISCUSSION

The results of one week from Sunday to Saturday with both training and validation corresponding to one day data are shown in Figure 3 to Figure 9. Generally, the results imply that the NARX is able to predict the consumer load as the predicted data is close to the actual with average MSE value of 0.0379.

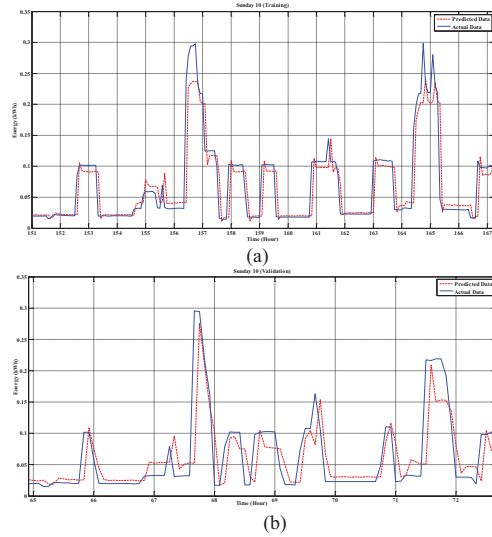


Figure 3: Day 10 Load Prediction (a) Training (b) Validation

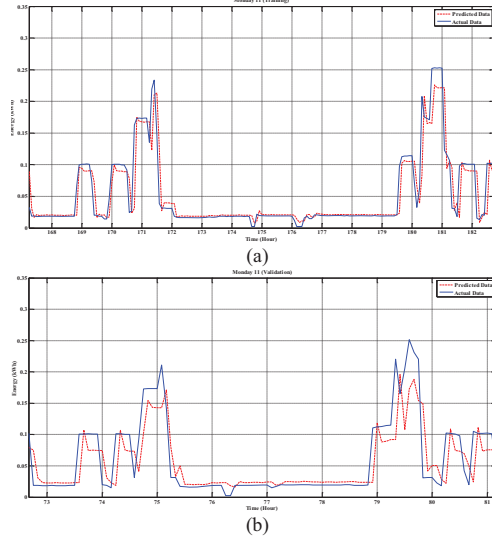
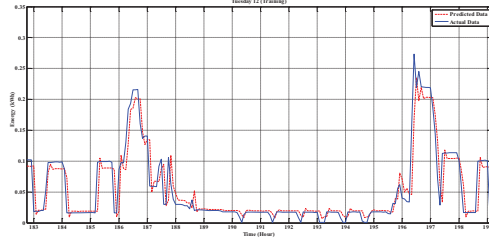
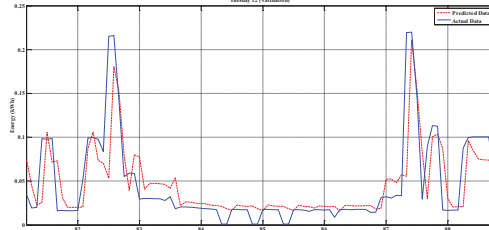


Figure 4: Day 11 Load Prediction (a) Training (b) Validation

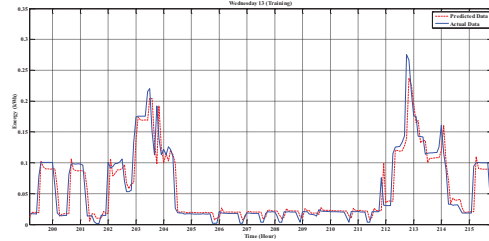


(a)

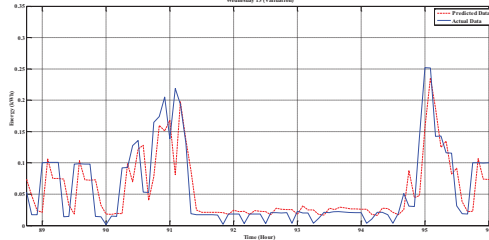


(b)

Figure 5: Day 12 Load Prediction (a) Training (b)Validation

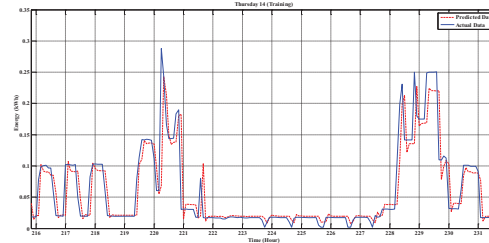


(a)

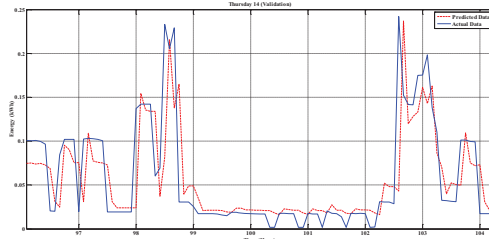


(b)

Figure 6: Day 13 Load Prediction (a) Training (b)Validation

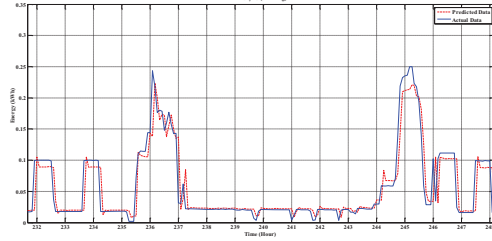


(a)

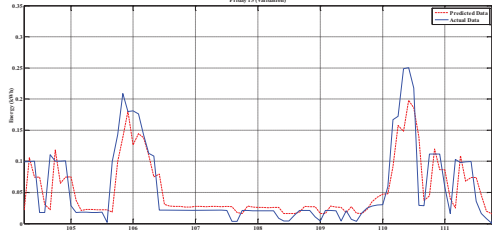


(b)

Figure 7: Day 14 Load Prediction (a) Training (b)Validation

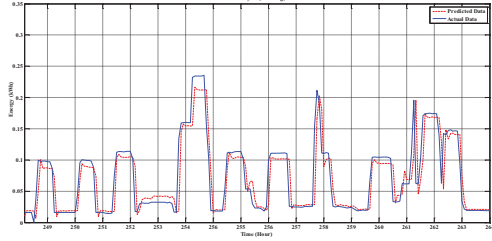


(a)

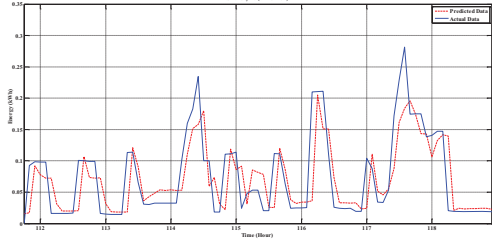


(b)

Figure 8: Day 15 Load Prediction (a) Training (b)Validation



(a)



(b)

Figure 9: Day 16 Load Prediction (a) Training (b)Validation

VII. CONCLUSION

The use of NARX network to predict consumer load has been thoroughly examined in this paper. The NARX architecture involves two tapped delay and nine hidden neuron. Average energy consumption data acquired from consumer load prototype for one month, with 288 data points per day was used and 5-minute step ahead load prediction was achieved. This facilitates the monitoring of individual consumer activities connected to the power distribution network and can be used to detect electricity theft when there is a huge disparity in the actual and predicted load. Although the prediction was okay however, the complexity of the NARX network is of concern because it has a total of 9 hidden neurons as well as two input and output tapped delay. Besides, the tapped delays were selected by trial and error which will be difficult to tune if applied in real time. Hence, there is a

need to modify the NARX optimally. This will be the focus of our next work.

REFERENCES

- [1] I. Koprinska, M. Rana, and V. Agelidis, "Electricity Load Forecasting: A Weekday-Based Approach," in *Artificial Neural Networks and Machine Learning – ICANN 2012*, vol. 7553, A. Villa, W. Duch, P. Érdi, F. Masulli, and G. Palm, Eds., ed: Springer Berlin Heidelberg, 2012, pp. 33-41.
- [2] A. Ghanbari, E. Hadavandi, and S. Abbasian-Naghneh, "An Intelligent ACO-SA Approach for Short Term Electricity Load Prediction," in *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence*, vol. 6216, D.-S. Huang, X. Zhang, C. Reyes García, and L. Zhang, Eds., ed: Springer Berlin Heidelberg, 2010, pp. 623-633.
- [3] H. Chen, K. Huang, and L. Chang, "Application of neural networks for very short-term load forecasting in power systems," in *Advances in Neural Networks-ISNN 2005*, W. Jun, L. Xiao-Feng, and Y. Zhang, Eds., ed Berlin Heidelberg: Springer, 2005, pp. 628-633.
- [4] W. Charytoniuk and M.-S. Chen, "Very short-term load forecasting using artificial neural networks," *Power Systems, IEEE Transactions on*, vol. 15, pp. 263-268, 2000.
- [5] J. V. Paatero and P. D. Lund, "A model for generating household electricity load profiles," *International journal of energy research*, vol. 30, pp. 273-290, 2006.
- [6] A. I. Abdullateef, M. J. E. Salami, I. B. Tijani, and M. A. Onasanya, "Novel technique for detecting electricity theft on low voltage distribution network," in *Proceeding of National conference of electrical and electronics engineering*, Johor, 2012, pp. 199-204.
- [7] A. Nizar and Z. Dong, "Identification and detection of electricity customer behaviour irregularities," Seattle, WA, 2009, pp. 1-10.
- [8] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, pp. 23-34, 2002.
- [9] J. W. Taylor, "An evaluation of methods for very short-term load forecasting using minute-by-minute British data," *International Journal of Forecasting*, vol. 24, pp. 645-658, 2008.
- [10] R. P. Singh, P. X. Gao, and D. J. Lizotte, "On hourly home peak load prediction," in *Third International Conference on Smart Grid Communications (SmartGridComm)*, Tainan, 2012, pp. 163-168.
- [11] I. Koprinska, M. Rana, and V. G. Agelidis, "Electricity load forecasting: a weekday-based approach," in *Artificial Neural Networks and Machine Learning-ICANN 2012*, E. V. Alessandro, D. Włodzisław, É. Péter, M. Francesco, and P. Günther, Eds., ed Berlin Heidelberg: Springer, 2012, pp. 33-41.
- [12] D. J. Trudnowski, W. L. McReynolds, and J. M. Johnson, "Real-time very short-term load prediction for power-system automatic generation control," *Control Systems Technology, IEEE Transactions on*, vol. 9, pp. 254-260, 2001.
- [13] C. O. Adika and L. Wang, "Short term energy consumption prediction using bio-inspired fuzzy systems," in *North American Power Symposium (NAPS)*, Champaign, 2012, pp. 1-6.
- [14] W. Yan, "Toward automatic time-series forecasting using neural networks," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 23, pp. 1028-1039, 2012.
- [15] N. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey," *Computational Intelligence Magazine, IEEE*, vol. 4, pp. 24-38, 2009.
- [16] E. Diaconescu, "The use of NARX neural networks to predict chaotic time series," *WSEAS Transactions on Computer Research*, vol. 3, pp. 182-191, 2008.
- [17] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, "Very Short-Term Load Forecasting: Wavelet Neural Networks With Data Pre-Filtering," *Power Systems, IEEE Transactions on*, vol. 28, pp. 30 - 41 2013.
- [18] H. Y. Yang, H. Ye, G. Wang, J. Khan, and T. Hu, "Fuzzy neural very-short-term load forecasting based on chaotic dynamics reconstruction," *Chaos, Solitons & Fractals*, vol. 29, pp. 462-469, 2006.
- [19] C. Y. Fok and M. I. Vai, "Very Short Term Load Forecasting for Macau Power System," in *Intelligent Computing Technology*, H. De-Shuang, J. Changjun, B. Vitoantonio, and C. F. Juan, Eds., ed Berlin Heidelberg: Springer, 2012, pp. 538-546.
- [20] G. Zheng, Y. Gao, and L. Wang, "Realization of Automatic Meter Reading System Based on ZigBee with Improved Routing Protocol," in *Power and Energy Engineering Conference (APPEEC)*, Chengdu, 2010, pp. 1-6.
- [21] G. Fuxiang, X. Wenxin, and L. Langtao, "Overview on remote meter reading system based on GPRS," in *2nd International Conference on Industrial and Information Systems (IIS)*, Dalian, 2010, pp. 270-273.
- [22] H. G. R. Tan, C. Lee, and V. Mok, "Automatic power meter reading system using GSM network," in *International Power Engineering Conference (IPEC)*, Singapore, 2007, pp. 465-469.
- [23] L. Wang and V. Devabhaktuni, "Support vector machine based data classification for detection of electricity theft," in *Power Systems Conference and Exposition (PSC)*, Phoenix, AZ, 2011, pp. 1-8.
- [24] A. Pasdar and S. Mirzakhaki, "A solution to remote detecting of illegal electricity usage based on smart metering," in *2nd International Workshop on Soft Computing Applications (SOFA)*, Oradea, 2007, pp. 163-167.
- [25] C. Bandim, J. Alves Jr, A. Pinto Jr, F. Souza, M. Loureiro, C. Magalhaes, et al., "Identification of energy theft and tampered meters using a central observer meter: a mathematical approach," in *Transmission and Distribution Conference and Exposition*, Rio de Janeiro, 2003, pp. 163-168.
- [26] I. H. Cavdar, "A solution to remote detection of illegal electricity usage via power line communications," *Power Delivery, IEEE Transactions on*, vol. 19, pp. 1663-1667, 2004.
- [27] S. Naiman, M. Kissaka, and N. Mvungi, "Energy meter reading and tampering protection through powerline communication channel," in *7th AFRICON Conference in Africa*, 2004, pp. 821-826.
- [28] B. Bat-Erdene, S. Y. Nam, and D. H. Kim, "A Novel Remote Detection Method of Illegal Electricity Usage Based on Smart Resistance," in *Future Information Technology*, J. J. Park, L. T. Yang, and C. Lee, Eds., ed Berlin Heidelberg: Springer, 2011, pp. 214-223.
- [29] P. Kadurek, J. Blom, J. Cobben, and W. Kling, "Theft detection and smart metering practices and expectations in the Netherlands," in *Innovative Smart Grid Technologies Conference Europe (ISGT)*, Gothenburg 2010, pp. 1-6.
- [30] A. Andalib and F. Atry, "Multi-step ahead forecasts for electricity prices using NARX: a new approach, a critical analysis of one-step ahead forecasts," *Energy Conversion and Management*, vol. 50, pp. 739-747, 2009.
- [31] Y.-W. Lee and T.-L. Chang, "Application of NARX neural networks in thermal dynamics identification of a pulsating heat pipe," *Energy Conversion and Management*, vol. 50, pp. 1069-1078, 2009.
- [32] S. Çoruh, F. Geyikçi, E. Kılıç, and U. Çoruh, "The use of NARX neural network for modeling of adsorption of zinc ions using activated almond shell as a potential biosorbent," *Bioresource Technology*, 2013.
- [33] L. Tsungnan, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *Neural Networks, IEEE Transactions on*, vol. 7, pp. 1329-1338, 1996.
- [34] Y. Gao and M. J. Er, "NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches," *Fuzzy sets and systems*, vol. 150, pp. 331-350, 2005.
- [35] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent NARX neural networks," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 27, pp. 208-215, 1997.
- [36] T. Lin, B. Horne, C. Giles, and S. Kung, "What to remember: how memory order affects the performance of NARX neural networks," in *International Joint Conference on Neural Networks & World*

Congress on Computational Intelligence Anchorage, AK, 1998, pp. 1051-1056.

- [37] A. I. Abdullateef, M. J. E. Salami, M. A. Musse, M. A. Onasanya, and M. I. Alebiosu, "New consumer load prototype for electricity theft monitoring," *IOP Conference Series: Materials Science and Engineering*, vol. 53, p. 012061, 2013.