Water Advisory Demand Evaluation and Resource Toolkit

Daniel Paluszczyszyn*[1], Sunday Iliya[2], Eric Goodyer[3], and Tomasz Kubrycht[4]

[1,2,3,4]De Montfort University, The Gateway, Leicester, LE1 9BH, UK

* *paluszcol@dmu.ac.uk*

ABSTRACT

The purpose of this feasibility study is to determine if the application of computational intelligence can be used to analyse the apparently unrelated data sources (social media, grid usage, traffic/transportation and weather) to produce credible predictions for water demand. For this purpose the artificial neural networks were employed to demonstrate on datasets localised to Leicester city in United Kingdom that viable predictions can be obtained with use of data derived from the expanding Internet-of-Things ecosystem. The outcomes from the initial study are promising as the water demand can be predicted with accuracy of 0.346 m^3 in terms of root mean square error.

Keywords: demand, prediction, computational intelligence

1 BACKGROUND

Cities are living organisms, 24h / 7day, with demands on resources and outputs. In particular, drinking water is the most valuable asset hence it is crucial to monitor, control and manage this resource. To do so water distribution system (WDS) operators need accurate water demand forecast for controlling the production, storage and delivery of drinking water. Despite that water is a key resource its management has not kept pace with modern urban life as demand for clean water and loads on waste water no longer fit diurnal patterns; and they are impacted by events that are outside the normal range of parameters that are taken account of in water management. The diurnal variation in the population compositions of urban areas is usually neglected in the traditional models of urban social structure, derived usually from standard census data [1]. As the water usage is linked to human activity, the ability to monitor or predict the population density fluctuations, and collective activity, can be used to provide advance control data for water management systems; for both the delivery of clean water and the removal of foul water.

This study, carried out within the scope of the Water Advisory Demand Evaluation and Resource (WADER) project, aims to determine if it is feasible to predict water demand in urban areas, particularly during severe weather events (excess rainfall and drought), by analysing e.g. social media usage, transport and meteorological data. Nowadays, much of these data can be derived from the emerging Internet-of-Things (IoT) ecosystem unleashing potential of data mining techniques in numerous applications including more tailored and efficient delivery of water to end users.

The WADER toolkit prediction engine uses computational intelligence (CI) techniques to analyse a mix of data inputs to produce credible predictions for clean water demand in urban areas. Especially, artificial neural networks, a subset of CI domain, are successfully used in the short and long term water demand predictions studies due to their inherent ability to detect patterns [2, 3, 4]; e.g. [2] reported that the artificial neural network (ANN) models for their short-term municipal water demand forecasting study consistently outperformed the regression and time-series models.

The final WADER toolkit data inputs will be social media activity, gas and electricity usage, combined with meteorological and traffic movement data. Such dataset should capture population fluctuations and activity over a subsequent prediction period, thus providing inputs to the water supply services on the future demand of fresh water supplies, and the subsequent load on waste water and sewerage systems. The computes the predictions in an open-source manner to support inter-operability; thus enabling the development of new applications.

Section 2 describes the WADER project concept, used approach, tools and built-in features. Section 3 discusses results from a preliminary case study. Section 4 concludes this paper.

2 WATER ADVISORY DEMAND EVALUATION AND RESOURCE TOOLKIT

2.1 Concept

The concept of the WADER project is to use traditional data sources (temperature, water and energy) that are augmented with social media and traffic flow data to deliver a more accurate predictions for the delivery of clean water, and removal of foul water from out cities. Social media usage data, coupled with electricity usage can indicate real-time fluctuations in the population density [5]. Traffic congestion data can be used to predict near-future load demand as people travel to work, home or places of recreation and commerce. Meteorological data derived can be used to predict weather patterns that will result in rain-water impacting on both water catchment areas, and drainage systems. These and other data sources will be used as the inputs to the WADER toolkit. Computational intelligence is used to drive WADER, with the CI system being trained using historical data, and predicted demand levels if demand for clean and waste water being presented using real-time analysis of live data feeds. The analysis of an aggregation of apparently unrelated data sources (social media, grid usage, traffic/transportation and weather) is a unique concept, that can be achieved by the deployment of computational intelligence.

2.2 Development and tools

The tool was created in the MATLAB software and employed capabilities its specialised toolboxes [6]. MATLAB provided means for development of graphical user interface (GUI), obtaining data from online database and creation of the prediction models. The developed MATLAB code interacts with the online database deployed with use of the Microsoft Azure technology [7]. Database management system was set up to collect data from a number of different sources including "live" data feeds e.g. from traffic flow observers.

The created tool can be used to develop various topologies of ANN models. It includes a sensitivity test feature to evaluate the importance or contribution of each of the input variable on the prediction accuracy of the model, and also as a means of comparing our approach with traditional methods of population and water prediction. The tool aims to provide predictions for different time intervals, e.g. hourly, daily, monthly and yearly. Embedded within the toolkit are variants of differential evolutionary and swarm intelligence optimisation algorithms for optimising the meta-parameters of the CI models e.g. the number of neurons and weights of ANN models.

2.3 Computational intelligence engine

Prediction models based on artificial neural network and support vector machine framework are used in this research for water demand prediction. ANN are massive parallel computing systems composed of large number of simple processing elements (processors) with many interconnections operating in parallel. These elements are inspired by the structure and function of biological nervous systems such as the brain [8]. The information processing units are called neurons which are interconnected together via a synaptic weights, and working together in parallel and unison to solve specific problems.

The activation functions play a vital role in ANNs, unfortunately optimum choice of activation functions for each neuron in a given network are problem dependent and can not be generalised. Due to the dynamic nonlinearity often associated with utilities consumption (e.g. electricity and gas) and how they impact on water demand, coupled with random weather variation, resulting from both artificial and natural sources, a fully connected multilayer perceptron (MLP) ANN with two hidden layers was used in this study. The input layer was cast into a high dimensional first hidden layer for proper features selection. In order to introduce a nonlinear transformation into the network, nonlinear hyperbolic tangent functions are used as the activation functions of the two hidden layers while a linear symmetric straight line is used for the output activation function. Other activation functions were also used, but this combination gave better promising results. The hidden layers are used to learn the salient features that characterizes the training data i.e. they serve as a features detector. The ANN is trained by combining the global search advantages of multiple solution metaheuristics algorithms and the local search advantages of single solution backpropagation algorithm (BPA) to further fine tune the weights toward the global optimum. The empirical risk function to be minimized is the mean square error (MSE).

A supervised batch training method was used with 60% of the data used for training the ANN, 20% for validation and 20% for testing ANN [9]. In this study, the back propagation algorithm is used as a local searcher, thus the learning rate was kept low at 0.01. Four ANN topologies were examined: feed-forward, cascaded feed-forward, feed-forward with output feedback and layered recurrent ANN.

2.4 Optimisation

A hybrid algorithm consisting of population based metaheuristic algorithm variants and single solution BPA is used to train the ANN model. The population based metaheuristics optimization algorithms used are based on the differential evolution [10] and swarm intelligence [11, 12] frameworks. The metaheuristic algorithms are used as global searchers to explore in detail the search space for evolving the initial weights and biases of the ANN model; after which the weights of the best candidate solution (model) are fine tuned using BPA towards the global optimum to produce the final optimised model. The BPA is used as a local searcher to exploit the region already explored by the metaheuristic, hence its learning rate and the momentum are set at low values (0.01 and 0.008 respectively). Despite the use of momentum constant and varying of the learning rate, BPA can easily be trapped in local optimum when solving multimodal problems, leading to premature convergence. To minimise this problem, the advantages of population based metaheuristic optimization algorithms is combined with the advantages of single solution BPA to evolve the weights of the ANN model.

Another challenge often encounter when designing ANN prediction models, is the choice of the number of neurons in each layer for near optimum performance. In order to circumvent this problem, the number of neurons in each layer were also evolved using the same metaheuristics. More details about the used optimisation algorithms and their settings can be found in [13].

2.5 Sensitivity test

The sensitivity analysis is feature of the WADER toolkit to test the contribution of each of the input or combination of the inputs with reference to a particular model using seven error measurements, i.e. MSE, root mean square error (RMSE), average absolute error (AAE), average relative percentage error (ARPE), discrepancy, long time volume error and long time volume percentage error. The feature aims to help in selecting the most relevant data for training and prediction for different scenarios. The sensitivity test evaluates the effect of the autocorrelation and cross-correlations of the inputs on the model prediction accuracy and generalisation. Since different models created (trained) using the same data are likely to capture different trends or patterns associated with the training data, to draw a conclusion on the significance of the inputs, the sensitivity test has to be conducted using different topologies (models) for a given number of times. The statistic of the error measurements can be used to evaluate the contribution of the inputs. Different error measurements are used to enable customers select the most appropriate inputs/models that best meet their target e.g. some customers may not be interested in discrete time accuracy but rather in long time volume error and so forth. This is particularly useful for selection of inputs from available data feeds; the sensitivity results may allow the costumer to use a model with inputs from free-of-charge data feeds while retaining accuracy and reducing the cost of the application. The error measurement to be used depend on the impact of the error on the level (cost) of risk.

For the illustration of sensitivity analysis benefits the WADER toolkit was set up to use an ANN model with two hidden layers. However, while the accuracy of ANN two-layer model is sufficient for demonstration of sensitivity analysis feature, it is recommended that the actual sensitivity analysis should use ANN models with more hidden layers.

The results from the sensitivity analysis are listed in Tables 1 and 2; Table 1 contains results for models trained with a past water usage where models in Table 2 were trained with water data. Comparing the results from the both tables it can be observed that for the ANN models used in the sensitivity test (only two hidden layers) the models without use of water input data would perform poorer. From Table 1 it may be concluded that the ANN model, which inputs are time and gas, performs the best in terms of average MSE (0.67089), average RMSE (0.8179) or average AAE (0.5528), while the ANN model, which inputs are time, temperature and electricity performs the best in terms of average Vol Err (140). These results may be used produce a model which meets requirements of a specific scenario. For example, if user is focused on short-term prediction (hours) and high accuracy at discrete time intervals should use the ANN model which inputs are time and gas only. In contrast, the user with focus on long-time prediction (days, months) should use the ANN model with time, temperature and electricity as it would achieve more accurate total volume prediction over the period of time. Further suggestion is to develop a weighted metric, which would reflect user needs in trade-off between the utilised error measurements.

When the inputs for the ANN are determined, the sensitivity test would enable selection the best model amongst tested models. Note that values in Tables 1 and 2 are average for the particular criteria. For example, from Table 1 one can choose the ANN model with time and gas as aforementioned. Figure 1 illustrates performance of particular models from the test. In this case (with time and gas inputs) five ANN models with different initial parameters were evaluated.

Table 1: The results from the sensitivity analysis carried out for the ANN model with use of the water input. MSE - mean square error (m^6), RMSE - root mean square error (m^3), AAE - absolute average error (m^3), ARPE - average relative percentage error, Vol Err - total volume error (m^3), Vol % Err - total volume percentage error, AMV - actual mean volume (m^3/h), PMV - predicted mean volume (m^3/h), ATV - actual total volume, PTV - Predicted total volume, SD - standard deviation.

Inputs	MSE		RMSE		AAE		ARPE %		Vol Err		Vol % Err		Discrepancy		AMV	PMV	ATV	PTV
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	Mean		
T	0.7403	0.0927	0.8591	0.537	0.5863	0.0588	35.26	4.29	270	57.53	5.2	1.16	0.33	0.44	2.6	2.68	5213.6	5360.4
T, G	0.6708	0.0804	0.8179	0.047	0.5528	0.0468	33.22	4.6	198	136	3.85	2.63	0.51	0.17	2.58	2.65	5154.6	5300.4
T, Te	1.0107	0.4674	0.9856	0.221	0.6901	0.1917	46.18	22.9	297	337	5.74	6.53	0.83	0.81	2.58	2.67	5164	5356.4
T, E	0.7234	0.0628	0.8499	0.036	0.5858	0.0481	35.44	5.79	192	209	3.7	4.05	0.51	0.34	2.62	2.7	5232.1	5404.3
T, G, Te	0.7723	0.3001	0.8678	0.154	0.5986	0.1442	38	14.9	266	253	5.18	5.01	0.63	0.49	2.6	2.74	5206.6	5472.8
T, G, E	0.7801	0.0647	0.8827	0.036	0.6191	0.0442	36.98	3.34	205	115	3.96	2.22	0.72	0.8	2.6	2.63	5203	5260
T, Te, E	0.83	0.066	0.9105	0.036	0.6288	0.0356	36.23	3.22	140	144	2.69	2.75	0.45	0.39	2.61	2.56	5215	5123
T, G, Te, E	0.8119	0.0842	0.9	0.048	0.6558	0.0456	41.57	6.25	288	275	5.59	5.37	0.67	0.38	2.6	2.74	5190	5470

Inputs: time (T)(implies only time variables (i.e. Min, Hour, day(1-7), month(1-12) and year) are used as inputs without any utility), gas (G), temperature (Te), electricity (E)

Table 2: The results from the sensitivity analysis carried out for the ANN model without use of the water input.

Inputs	MSE		RMSE		AAE		ARPE %		Vol Err		Vol % E		Discrepancy		AMV	PMV	ATV	PTV
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	Mean		
T	2.113	0.0852	1.453	0.029	1.03	0.025	60.94	3.76	218	213	4.03	3.94	3.93	0.61	2.7	2.61	5413	5216
T, G	5.96	0.48	2.44	0.099	1.8	0.092	64.42	2.44	3330	363	61	6.71	6.91	0.44	2.71	1.04	5413	2082
T, Te	2.34	0.174	1.53	0.055	1.094	0.028	64.17	6.93	395	447	7.31	8.27	4.37	0.6	2.71	2.51	5413	5027
T, E	2.98	0.3221	1.73	0.092	1.298	0.0866	83.47	7.41	499	283	9.23	5.23	4.84	0.706	2.71	2.46	5413	4914
T, G, Te	5.92	1.2043	2.42	0.265	1.79	0.199	69.05	5.07	3011	964	55.63	17.8	7.38	0.968	2.71	1.2	5413	2402
T, G, E	5.5	1.247	2.33	0.271	1.776	0.258	78.99	30	2747	1096	50.76	20.3	6.76	0.96	2.71	2.13	5413	4254
T, Te, E	4.43	1.0415	2.09	0.243	1.599	0.214	93.3	15.93	1205	695	22.26	12.85	4.89	1.108	2.71	1.16	5413	4256
T, G, Te, E	5.71	1.1076	2.38	0.234	1.816	0.196	74.17	9.49	3097	648	57.22	11.99	6.91	0.77	2.71	1.16	5413	2315

Inputs: time (T)(implies only time variables (i.e. Min, Hour, day(1-7), month(1-12) and year) are used as inputs without any utility), gas (G), temperature (Te), electricity (E)

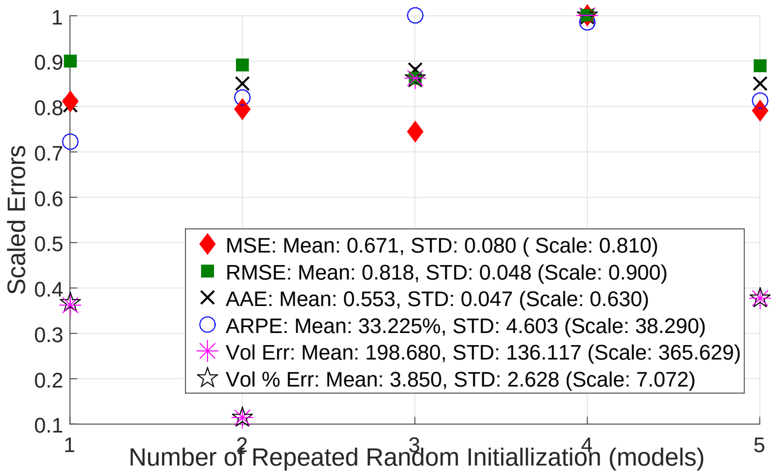

Figure 1: Illustrating particular models performance for different random initialisation parameters. Note that the errors were normalised for clearer illustration. To retrieve actual value one must use the scale values given in the legend.

Similar to the above analysis; if costumer is interested in small MSE and RMSE metrics it is recommended to use model 3 (indicated by number 3 on horizontal axis), if costumer is more interested in predicting more accurately total water volume over a period of time the model 2 is recommended. It important to highlight that the sensitivity test is not a proof of the accuracy of the final model. It is rather a mean of determining the most likely inputs to be used for the design of the final model. It is therefore needful that the test should be conducted on wide array of models including poor performance models in order to observe the impact of the inputs on the prediction accuracy. The test will also help to determine the necessary inputs (those that must be available) in addition to other inputs in order to obtain a given performance index.

3 PRELIMINARY RESULTS

To test the functionality of the developed tool along with appropriateness of the proposed approach, the data obtained from the SmartSpaces project website http://smartspaces.dmu.ac.uk were utilised. The SmartSpaces project monitors the energy performance of a selection of public buildings in Leicester, UK (university buildings, city council buildings, schools, libraries, leisure centres and others). The data are collected at 30 min intervals and include ambient temperature, electricity, gas and water usage (At this stage of the WADER project data from traffic and social media were not yet available).

Performance of three different ANN models, depicted in Figure 2, demonstrate functionality of the toolkit, which allows user to gradually improve the prediction model either by employing additional data or the optimisation metaheuristic. The top plot illustrates the ANN model prediction performance with no past water data input to the model. The middle plot depicts the prediction model which includes the water usage. The model in the bottom plot had the same initial parameters for training with the latter model but its number of neurons and weights were subsequently optimised. The initial MSE error of 1.285 (top plot) was reduced to 0.759 (bottom plot). With the data obtained from the SmartSpace project the toolkit is able to obtain models with the prediction accuracy below 0.2 in terms of MSE as illustrated in Figure 3. However, what was interesting is that models with no water feed can still predict water usage fairly accurate (see top plot in Figure 2). This might be particularly useful in

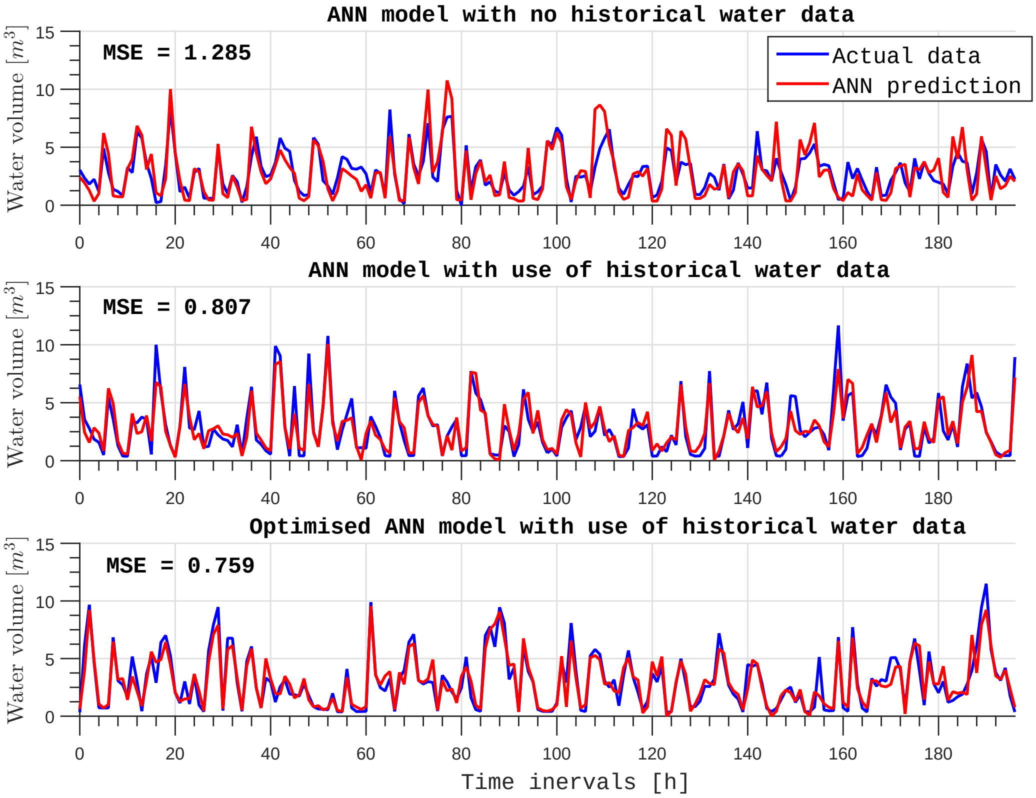

Figure 2: Performance of different ANN prediction models.

situations when the water data feeds are not available.

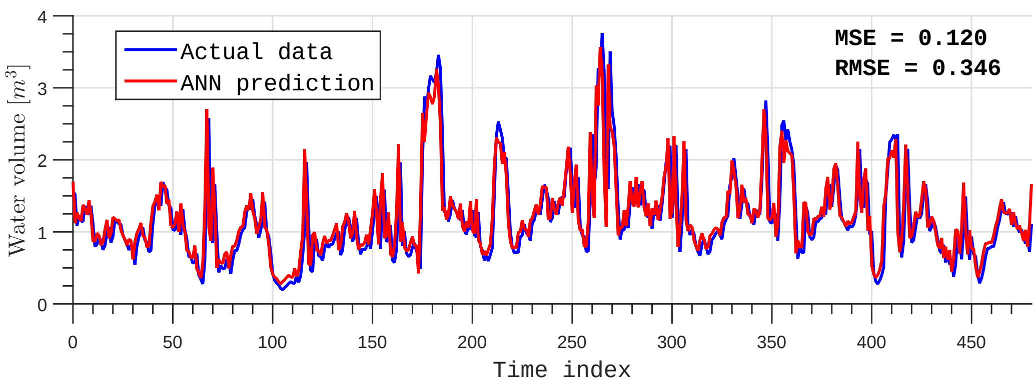

Figure 3: Illustrating the accuracy of ANN prediction models.

4 CONCLUSIONS

The WADER project employs ANN-based models that uses learning of past usage, to predict future water demand. In urban areas, the drinking water usage is closely associated with daily population distribution, which can be monitored with use of electricity, gas, social media and traffic to enhance water demand prediction. The preliminary study with only gas, electricity and ambient temperature data demonstrated that water demand can be predicted with a sufficient accuracy. Additional data sourced from social media and traffic could not only improve fidelity of the developed prediction models but also result in a different application of the WADER toolkit; i.e. prediction of daily fluctuations in population density in urban areas.

References

[1] M. F. Goodchild and D. G. Janelle, "The city around the clock: Space - time patterns of urban ecological structure," *Environment and Planning A*, vol. 16, no. 6, pp. 807–820, 1984.

[2] J. Bougadis, K. Adamowski, and R. Diduch, "Short-term municipal water demand forecasting," *Hydrological Processes*, vol. 19, no. 1, pp. 137–148, 2005.

[3] M. Ghiassi, D. K. Zimbra, and H. Saidane, "Urban water demand forecasting with a dynamic artificial neural network model," *Journal of Water Resources Planning and Management*, vol. 134, no. 2, pp. 138–146, 2008.

[4] M. Romano and Z. Kapelan, "Adaptive water demand forecasting for near real-time management of smart water distribution systems," *Environmental Modelling & Software*, vol. 60, pp. 265 – 276, 2014.

[5] P. A. Grabowicz, J. J. Ramasco, B. Gonçalves, and V. M. Eguíluz, "Entangling mobility and interactions in social media," *PloS one*, vol. 9, no. 3, p. e92196, 2014.

[6] Mathworks. (2016) Matlab - the language of technical computing. [Online]. Available: http://uk.mathworks.com/products/matlab

[7] Microsoft, "Microsoft azure: Cloud computing platform and services," 2017. [Online]. Available: https://azure.microsoft.com/

[8] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Pearson Education, Inc., Upper Saddle River, New Jersey 07458, 2008.

[9] A. Abbasi, R. Y. K. Lau, and D. E. Brown, "Predicting behavior," *IEEE Intelligent Systems*, vol. 30, no. 3, pp. 35–43, 2015.

[10] K. V. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization.* Springer, 2005.

[11] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Evolutionary Computation Proceedings, IEEE World Congress on Computational Intelligence.* IEEE, 1998, pp. 69–73.

[12] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 281–295, 2006.

[13] S. Iliya, "Application of computational intelligence in cognitive radio network for efficient spectrum utilization, and speech therapy," Ph.D. dissertation, De Montfort University, 2016.