AAHES: A Hybrid Expert System Realization of Adaptive Autonomy for Smart Grid


Abstract—Smart grid expectations objectify the need for optimizing power distribution systems greater than ever. Distribution Automation (DA) is an integral part of the SG solution; however, disregarding human factors in the DA systems can make it more problematic than beneficial. As a consequence, Human-Automation Interaction (HAI) theories can be employed to optimize the DA systems in a human-centered manner. Earlier we introduced a novel framework for the realization of Adaptive Autonomy (AA) concept in the power distribution network using expert systems. This research presents a hybrid expert system for the realization of AA, using both Artificial Neural Networks (ANN) and Logistic Regression (LR) models, referred to as AAHES, respectively. AAHES uses neural networks and logistic regression as an expert system inference engine. This system fuses LR and ANN models’ outputs which will result in a progress, comparing to both individual models. The practical list of environmental conditions and superior experts’ judgments are used as the expert systems database. Since training samples will affect the expert systems performance, the AAHES is implemented using six different training sets. Finally, the results are interpreted in order to find the best training set. As revealed by the results, the presented AAHES can effectively determine the proper level of automation for changing the performance shaping factors of the HAI systems in the smart grid environment.


I. NOMENCLATURE

AAHES: Adaptive Autonomy using Hybrid Expert System
DA: Distribution Automation
HAI: Human-Automation Interaction
AA: Adaptive Autonomy

A. Fereidunian is with Islamic Azad University, Islamshahr Branch, Tehran, IRAN (fereidunian@iiau.ac.ir & arf@ece.ut.ac.ir).
M.A. Zamani is with CIPCE, School of ECE, University College of Engg., University of Tehran, Tehran, IRAN (maz.zamani@gmail.com & ma.zamani@ece.ut.ac.ir).
F. Boroomand is with ECE Dep., Concordia University, Montreal, QC, CANADA (F boroomand@ece.ut.ac.ir).
R. Jamalabadi is with CIPCE, School of ECE, University College of Engg., University of Tehran, Tehran, IRAN (h.jamalabadi@ece.ut.ac.ir).
H. Lesani is with CIPCE, School of ECE, University College of Engg., University of Tehran, Tehran, IRAN (lesani@ut.ac.ir).
C. Lucas is with CIPCE, School of ECE, University College of Engg., University of Tehran, Tehran, IRAN.
Sh. Shariat-Torbaghan is with Dep. of Energy and Environment, Chalmers University of Technology, SWEDEN (sshahab@student.chalmers.se).
M. Meydani is with CIPCE, School of ECE, University College of Engg., University of Tehran, Tehran, IRAN (mohammadmeydani@gmail.com).

LR: Logistic Regression
ANN: Artificial Neural Network
LOA: Level of Automation
CCR: Correct Classification Rate
GTEDC: Greater Tehran Electric Distribution Company
PSF: Performance Shaping Factors
UMA: Utility Management Automation
SCADA: Supervisory Control and Data Acquisition
MLE: Maximum Likelihood Estimation

II. INTRODUCTION

MODERNIZATION of distribution grid is a core objective in smart grid innovations, and DA is an integral part of that. DA allows real-time monitoring, control and automated operation of distribution networks. It will also improve the efficiency and reliability of the power distribution networks [1], [2].

Alongside with all of these advantages, it should be considered that disregarding human factors in the automation systems makes them more problematic than beneficial [3]. In fact, the human-automation "team" is more productive than either human or automation working alone [4]. Considering both human factors and automation systems simultaneously leads to a high level of complexity. In order to manage this complexity, the HAI concept has been extensively studied in recent years. A simple form of the HAI model was first introduced by P.M. Fitts in 1951, where only two levels of automation (manual or automate) were considered [5]. Since this primary model was no longer successful in optimizing human-automation interaction, Sheridan and Verplank introduced a ten-degree level of automation (LOA), to overcome the deficiency of Fitts' two-degree model [6]-[10]. Afterwards, Parasuraman, Sheridan and Wickens [6] suggested the AA concept (also known as adaptive automation or adjustable automation [6], [8]-[11]), which is expected to adapt the LOAs to the environmental conditions; in order to optimize human-automation systems performance in different environmental conditions [6],[10]. Subsequently, Fereidunian et al introduced a model-based framework for realization of the AA concept [12]-[14], and suggested expert systems for human-automation complexity management [15]-[17].

Although considerable amount of research have been dedicated to this concept, still more investigations are required to implement the HAI and the AA concepts in industry and civil services [7]. Excluding military and aerospace applications, [13] and [14] report the first implementations of
the AA expert systems in the civil services. However, the capability of the simple model introduced in [13], [14] is partially acceptable in tracking and simulating human experts' opinion in complicated situations.

This article—as a continuum of a series—presents a hybrid expert system, using the ANN and the LR, which improves its ancestors’ characteristics ([15]-[17]), in terms of higher Correct Classification Rate (CCR) and more accurate predictions in complex situations. The expert system introduced in this paper (referred to as AAHES) is a decision-fusion combination of LR and ANN outputs, which provide progressive results, comparing to both models' outputs.

The reminder of the paper is organized as follows: the methodology of AAHES realization, followed by implementation and results of the system. Afterward, a discussion is presented, in order to investigate the performance of the proposed AAHES.

III. METHODOLOGY

This paper presents a hybrid expert system, using neural networks and logistic regression. The practical data for this research are obtained from Grater Tehran Electric Distribution Company (GTEDC).

A. Position of AAHES

Reference [11] introduced a framework for implementation of the AA concept, where it was suggested to extract environmental conditions and represent them in a binary vector for further calculations. The most effective parameters on human-automation system were extracted and named Performance Shaping Factors (PSFs) [13], [14]. The AAHES having a PSF vector can adapt the LOAs to the environmental conditions.

The Human-automation system studied in this research is electric utility management automation (UMA) system, which is a sort of Supervisory Control and Data Acquisition (SCADA) system. This is an integral part of a smart grid solution. Fig. 1 illustrates the position of the AAHES in the UMA system. As shown in Fig. 1, the AAHES takes the PSFs from the UMA system, and recommends the proper LOA for that specific PSF, as a closed loop control system.

B. AAHES Structure

We introduced the AALRES as an expert system that is realized by logistic regression [17]. AALRES confirmed that the LR model is pessimistic, which means that the output of the models is almost always less than the expected values. Therefore, we expect that the output of the AALRES can be improved by fusing its outputs (i.e. aggregating) with an optimistic model. According to our experience (at least, in this application), artificial neural networks (ANNs) [19] are optimistic estimators of the LOAs (See Fig. 2), compared to that of the LR. As a result, our proposed hybrid expert system takes advantages of two different models; based on input-output data, collected through interviews with human experts in the GTEDC.

Any change in environmental conditions transits the PSF vector to a new state; as a result, the expert system recalculates the appropriate LOA, which may be different from the previous one. The appropriate LOA is determined through both LR and ANN models which are pre-trained by a set of data.

The following hybrid structure is presented for AAHES: the same training set is employed for training both LR and ANN models. Considering that the LR and ANN outputs are confined to [0, 1], the output for training samples should be scaled to this interval. Consequently, the final output should be again rescaled to a [0, 10] interval, and rounded to the nearest integer in order to indicate one of the Sheridan's ten levels of automation (LOA).

C. ANN Training

The feed-forward multilayer network is selected for the ANN structure. The number of input nodes are equal to the number of the PSFs (here, ten input nodes are considered). Furthermore, 20 neurons are designated for the hidden layer, where the output of a single sigmoid neuron determines the appropriate LOA. Fig. 3 shows the network structure for the ANN model.

The transfer function employed in this modeling is sigmoid function which is similar to the logit function (it will be presented in the next section). Their similarities raise the assumption of generalizing some features of one model to the other.

D. LR Training

LR is another machine learning algorithm that is employed to implement this expert system (AAHES). The capability of the LR in dealing with multidimensional variables makes it suitable for our expert systems realization. Like previously discussed, the ANN model, the LR model receives the PSF vector as its input, and recommends ascalor number as its output that must be rounded to an integer, to indicate the LOAs. The vector \( \mathbf{X} = [x_1, x_2, \ldots, x_n] \) represents the PSFs. Where \( x_i \) represent the \( i^{th} \) PSF of the HAI system. Eqs. (1) and (2) show the relationship between the input vector (PSF) and the output scalar [20]:

\[
\text{PSF} = \mathbf{X} = [x_1, x_2, \ldots, x_n]
\]
where \( \beta_0 \) is the intercept, and \( \beta_1 \) to \( \beta_{10} \) are the regression coefficients of \( x_i \), which are determined by employing Maximum Likelihood Estimation (MLE). MLE is the generalization of the least mean square method for the nonlinear models [21], [22]. For a fixed set of data and underlying probability model, maximum likelihood picks \( \beta \)s so that the data will be more likely than to any other values among these parameters (\( \beta \)s).

E. Fusion algorithm

The fusion algorithm -which has been obtained through our experiment in this application-, is based on the following principles:

1. The LR model is more accurate in predicting high levels of automation.
2. The ANN model is more accurate in predicting low levels of automation.

It should be mentioned that these rules have been determined through observations on our training samples. We present the following fuzzy linguistic rules, based on the previous mentioned principles:

1. **If** the predicted LOA of the LR model is equal to or more than seven, **then** the AAHES output is equal to the LR one.
2. **If** the predicted LOA of the ANN model is equal to or less than three, **then** the AAHES output is equal to the ANN one.
3. **Else**, the AAHES output is the average of the LR and ANN models, which is finally rounded to its nearest integer in the \([0, 10]\) interval.

IV. IMPLEMENTATION AND RESULTS

A. Implementation

The HAI studied in this research is an electric UMA system, which is a sort of the SCADA system for the electric power utility system [23].

Once a data-driven model is selected, choosing an appropriate training set can effectively improve the intelligence of the systems [24], [25]. In order to fulfill the optimal training, different training sets are selected and taught to the expert system. Since the expert system is expected to simulate experts’ opinion, both training set and test set are compared to the superior experts’ opinion. The superior experts are experts, whose superiority (in higher and more reliable expertise) has been verified according to the consistency of their expert judgments interview questionnaire [15].

Different combinations of the PSFs, orchestrate 324 feasible states. Each PSF illustrates the existence or non-existence of a condition in human-automation interaction system; however, some of these conditions are impossible to occur simultaneously.

B. Scenario Development

We developed six different scenarios for implementing this expert system. These scenarios are realized for both ANN, and LR models. Finally the AAHES is implemented, based on the ANN, and the LR outputs. In the following, the six scenarios are developed and their relevant results are presented latter:

- **Scenario 1**: Selection of 60 simple to complex samples as the training set
- **Scenario 2**: Selection of 60 complex to simple samples as the training set
- **Scenario 3**: Selection of 60 random combinations of the PSFs as the training set
- **Scenario 4**: Selection of 100 simple to complex samples
- **Scenario 5**: Selection of 100 complex to simple samples as the training set
- **Scenario 6**: Selection of 100 random combinations the PSFs as the training set

C. Results for Scenarios

**Scenario 1: Selection of 60 simple to complex samples as the training set**

Sixty samples are selected within the all 324 feasible states, in order to train the AAHES. These samples are selected with an incremental order in terms of complexity of the training samples. The more complexity stands for more ones in binary PSF vectors. Since this training set is not that much complex, it will be partially successful in equipping the AAHES to deal with complex situations. However, this training set facilitates the human experts to judge on proper LOAs, and consequently, their judgments would become more reliable. It should be considered that to perform an evaluation we are obliged to enquire all the feasible samples in this particular application. Ultimately, this kind of training set may cause the
expert system to hardly percept the expertise. Because, it should generalize its prior knowledge to more complex situations, and recommend the proper LOA for the more complex PSF vectors (Here means more than four simultaneously occurring PSFs). Table I shows the results for the AAHES, LR, and ANN models.

| Table I. The results for AAHES, LR, and ANN models for Scenario 1 |
|---------------------------------|-----------------|-----------------|-----------------|
| Model                          | CCR train       | CCR test        | CCR total       |
| LR model                       | 80%             | 42%             | 49%             |
| ANN model                      | 100%            | 37%             | 48%             |
| Hybrid Model                   | 95%             | 49%             | 58%             |

Scenario 2: Selection of 60 complex to simple samples as the training set

In this scenario the most complex sixty PSF vectors are selected to train the AAHES. Since this training set describes more complex situations comparing to scenario 1, it is expected to provide the AAHES with more amount of embedded information inside. However, this complexity faces the human experts' judgments with difficulty; consequently, their judgments become less reliable. It should be considered that to perform an evaluation, we are obliged to ask all the feasible samples in this particular application. Comparing to the training set in scenario 1, this kind of training set improves the generalization ability of the expert system; in other words, it should deduce the less complex situations from its prior knowledge, and recommend the proper LOA for the less complex PSF vectors. Table II shows the results for the AAHES, LR, and ANN models.

<table>
<thead>
<tr>
<th>Table II. The results for AAHES, LR, and ANN models for scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>LR model</td>
</tr>
<tr>
<td>ANN model</td>
</tr>
<tr>
<td>Hybrid Model</td>
</tr>
</tbody>
</table>

Scenario 3: Selection of 60 random combinations of the PSFs as the training set

Sixty samples (out of 324) are selected randomly in order to train the AAHES. This training set is generally more complex compared to the training set examined in scenario 1; however less complex than the one presented in the scenario 2. Note that, this training set guarantees a uniform distribution of information over least to most complex situations. As a consequence, we expect that this training set improves the generalization ability of the AAHES. Although a random training set is the most suitable one to improve (the) generalization ability of the AAHES, it faces superior experts' judgments with more difficulties, because it will faces them with completely different situations during the interview process. Again it should be considered, to perform an evaluation we are obliged to ask all the feasible samples in this particular application. CCRs correspond to the AAHES, LR, and ANN models are reported in Table III.

| Table III. The results for AAHES, LR, and ANN models for scenario 3 |
|-----------------|-----------------|-----------------|
| Model           | CCR train       | CCR test        | CCR total       |
| LR model        | 90%             | 86%             | 87%             |
| ANN model       | 100%            | 64%             | 71%             |
| Hybrid Model    | 97%             | 86%             | 88%             |

Scenario 4: Selection of 100 simple to complex samples

The number of training samples needed for the training the LR model is ten times more than its inputs variables [21]. Moreover, the transfer function employed for the ANN model is a sigmoid function, which is similar to the logit (The link function for the LR model). This similarity provided us with the concept of increasing the AAHES training samples to 100 (Ten times more than the expert system inputs). One hundred samples are selected within all 324 feasible states in order to train the AAHES. These samples are selected with an incremental order in terms of complexity of the training samples. Increasing in the number of training samples makes human experts interviews somehow more time-consuming and difficult and this will lead to an increase in the human experts' mistakes when compared to the previous scenarios. Nevertheless, simple to complex training sets may cause the expert system hardly percept the expertise. This arises due to the fact that it should generalize its prior knowledge to more complex situations, and recommend the proper LOA for the more complex PSF vectors (Here means more than four occurring PSFs). Table IV shows the results for the AAHES, LR, and ANN models.

| Table IV. The results for AAHES, LR, and ANN models for scenario 4 |
|-----------------|-----------------|-----------------|
| Model           | CCR train       | CCR test        | CCR total       |
| LR model        | 82%             | 76%             | 78%             |
| ANN model       | 87%             | 60%             | 68%             |
| Hybrid Model    | 86%             | 83%             | 84%             |

Scenario 5: Selection of 100 complex to simple samples as the training set

In this scenario, 100 of the most complex PSF vectors are selected in order to train the AAHES. Since this training set describes more complex situations, it is expected to provide the AAHES with more amount of embedded information comparing to the scenario 4. However, this complexity will face the human experts' judgments with difficulty, and consequently their judgments would become less reliable. In comparison with the training set in the scenario 4, this kind of training set improves the generalization ability of the expert system, because it should deduce the less complex situations...
from its prior knowledge, and recommend the proper LOA for the less complex PSF vectors. Table V shows the results for the AAHES, LR, and ANN models.

| Scenario 6: Selection of 100 random combination PSFs as the training set |
|--------------------------|--------------------------|--------------------------|
| LR model | ANN model | Hybrid Model |
| CCR train | 96% | 99% | 97% |
| CCR test | 83% | 58% | 76% |
| CCR total | 87% | 71% | 82% |

TABLE V. The results for AAHES, LR, and ANN models for scenario 5

Another factor which affects the generalization ability of the expert system is the number of its training samples. According to [21], the necessary number of training samples for the LR model is ten times of the number of its input variables. According to the similarity between the link function of the LR model (logit), and the transfer function of the ANN model (Sigmoid), and also the similarity of the error functions which two models employ (least mean square), we suggest to generalize this fact for neural networks. This assumption has been confirmed by the simulation results. Fig. 4 and Fig. 5 show the effect of randomness and number of training samples on the AAHES.

In summary, the AAHES improves the performance of its ancestors in term of generalization ability in complex situations. It seems that this structure is suitable for simulating the complex systems, and situations.

V. DISCUSSIONS

The AAES that was presented in [13], was a successful step in calculating the proper LOA; however, in more complicated situations, some of the outputs of the AAES were not satisfactory in tracking the human experts' opinion, (i.e. working with more "ones" in the binary PSF vectors). Further, [17] implemented AALRES using LR, which expressed better results in tracking human experts' opinion in complicated the situations; however, the pessimistic behavior of the LR caused the AALRES to track (trace) the human experts' opinion a little downward. Per contra, this paper combines a logistic regression with artificial neural networks in order to compensate its pessimism with optimistic behavior of the neural networks which leads to a more realistic system.

Although the AAHES is more successful in tracking the human experts' opinion compared to its ancestors’, it needs more samples for training, since it uses the data-driven models. In the following, we devote our effort to interpret the results, in order to determine the proper training set which will lead to the maximum training performance.

The results illustrate that random training samples are more suitable for the both LR, and ANN models. This can be explained considering the fact that a random training set ensures an adequate amount of embedded information, which will improve the generalization ability of the AAHES.

VI. CONCLUSIONS

A hybrid expert system was introduced for realization of the AA framework of [13],[14], referred to as AAHES. The presented AAHES adapts the LOA of UMA (a part of the smart grid) to the environmental conditions. The judgments of GTEDC’s experts were developed as a subjective data-base for the AAHES.

Simulation results illustrated that the AAHES is successful in adapting the level of automation (LOA) to the environmental conditions of the automation system. It was also indicated that the generalization ability of the AAHES is
considerably affected by the method of training set selection. Finally, increasing the number of training samples by 66% (from 60 samples to 100 samples) improved total CCR of the system with 2%-5%. This fact popped up the assumption of highly correlated input variables in our mind. This research will be continued by defining effective number of input variables, and comparing AAHES to its alternatives.

VII. ACKNOWLEDGMENT

Many thanks to the SCADA experts in the Greater Tehran Electricity Distribution Company (GTEDC), for the time they afford for interviews and to fill in our questionnaires;

Special appreciation to GTEDC managing board, especially to: Mr Habib Torabi (Deputy CEO), Mr Payam Kharazmi (CIS/CRM Manager) and Mr A. Kazemabkhs (Technical and Operation Deputy) for their helpful practical assistance in preparing the experimental data for this research.

The authors would also thank to Mr Babak Eskandarnejad for his assistance in preparing this paper.

VIII. REFERENCES

[19] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 2004

BIOGRAPHIES

Alireza Fereidunian received his PhD and MSc from University of Tehran (School of ECE), in 2009 and 1997, where he is a Research Associate now. Dr. Fereidunian serves as an instructor at Islamic Azad University, Imamshahr Branch and as an independent consultant to GTEDC. His research interests include power systems automation, and application of intelligent systems, human-automation interaction, data mining, decision-support, decision-making, IT and signal processing in power systems. He is a member of IEEE (and IEEE-SMC-HCI TC member) and INCOSE (as INCOSE Iran point of contact).

Mohammad Ali Zamani is a Research Assistant at the Center of Excellence for Control and Intelligent Processing (CECIP), School of Electrical and Computer Engineering, University of Tehran, Iran. He received his B.Sc. from School of Electrical and Computer Engineering, University of Tehran, Iran. His research interests include System Engineering, Smart Grid, Artificial Intelligence and Adaptive Autonomy. He is currently working on the implementation of system engineering approaches in smart grid.

Farzam Boroomand is a graduate student at Concordia University, Montreal, QC, Canada. He received his B.Sc. from School of Electrical and Computer Engineering, University of Tehran. His research interests include System Engineering, Smart Grid, Artificial Intelligence and Adaptive Autonomy. His research interests include control systems, flight
control, artificial intelligence, power distribution system automation, decision-making, neural networks, fuzzy logic, and human-automation interaction.

**Hamid Lesani** is a Professor at the Center of Excellence for Control and Intelligent Processing (CIPCE), School of Electrical and Computer Engineering, University of Tehran. He received the M.S. degree in electrical power engineering from the University of Tehran, Iran, in 1975, and the Ph.D. degree in electrical engineering from the University of Dundee, U.K., in 1987. Early in his career, he served as a Faculty Member with Mazandaran University. After obtaining the Ph.D. degree, he joined the Department of Electrical and Computer Engineering, Faculty of Engineering, University of Tehran. His teaching and research interests are design and modeling of electrical machines and power systems. Professor Lesani is a member of IEEE (PES) and IEEE Iran Section.

**Caro Lucas** was a Professor and a member (as well as the Founder- Director) of Center of Excellence for Control and Intelligent Processing (CIPCE), School of Electrical and Computer Engineering, University of Tehran, as well as a Researcher at the School of Cognitive Sciences (SCS), Institute for Studies in Theoretical Physics and Mathematics (IPM). His research interests include uncertain systems, intelligent control, multi-agent systems, data mining, business intelligence, financial modeling, biological computing, computational intelligence, neural networks, and knowledge management. He was the founder of the ISRF, IPM and has assisted in founding several new research organizations and engineering disciplines in Iran. He is the holder of the patent for Speaker Independent Farsi Isolated Word Neurorecognizer. Dr. Lucas has served as Chairman of the IEEE, Iran Section (1990-1992) and as the Chairman of several international conferences. Professor Lucas has been nominated as a member of Long-lasting Figures of Science and Culture Hall of Fame by the Iranian Academy of Engineering in 2006. Regretfully, the late Professor Caro Lucas passed away on July 8, 2010.

**Shahob Shariat-Torbaghan** is a graduate student at Chalmers University of Technology, Sweden. He received his B.Sc. degree from University of Tehran, Iran in 2008. His research interests include power system management and control, demand response analysis, Smart Grids and power distribution system automation.

**Mohammad Meydani** is a B.Sc. Student at the Center of Excellence for Control and Intelligent Processing (CIPCE), School of Electrical and Computer Engineering, University of Tehran. His research interests include control and optimization of power systems, power distribution system automation, and load forecasting in distribution systems.