

# Development and Application of An Enhanced ART-Based Neural Network

Keem Siah Yap<sup>1</sup>, Chee Peng Lim<sup>2</sup>, Eric W.M Lee<sup>3</sup>, Junita Mohamed Saleh<sup>4</sup>

<sup>1</sup>Department of Electronics & Communication Engineering, Universiti Tenga Nasional, Malaysia

<sup>2</sup>School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Malaysia

<sup>3</sup>Department of Building and Construction, City University of Hong Kong, Hong Kong, China

<sup>4</sup>School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Malaysia

**Abstract-** The Generalized Adaptive Resonance Theory (GART) neural network is developed based on an integration of Gaussian ARTMAP and the Generalized Regression Neural Network. As in our previous work [13], GART is capable of online learning and is effective in tackling both classification and regression tasks. In this paper, we further propose an Ordered-Enhanced GART (EGART) network with pruning and rule extraction capabilities. The new network, known as O-EGART-PR, is equipped with an ordering algorithm that determines the sequences of training samples, a Laplacian function, a new vigilance function, a new match-tracking mechanism, and a rule extraction procedure. The applicability of O-EGART-PR to pattern classification and rule extraction problems is evaluated with a problem in fire dynamics, i.e., to predict the occurrences of flashover in a compartment fire. The outcomes demonstrate that O-EGART-PR outperforms other networks and produces meaningful rules from data samples.

**Keywords-** Adaptive Resonance Theory, Generalized Regression Neural Network, rule extraction, fire safety engineering

## I. INTRODUCTION

Over the last two decades, artificial neural networks (or simply *neural networks*) have been developed for solving many pattern classification problems [1–4]. In the medical domain, neural networks have been applied as a diagnostic decision support system, e.g., a supervised learning neural network was devised for leukemia diagnosis [5]. In [6], an incremental learning fuzzy neural network that is optimized by a genetic algorithm was developed for breast cancer diagnosis and detection. Another breast cancer detection and diagnosis approach using a combined numerical and linguistic system was presented in [7]. In [8] a hybrid model of Adaptive Resonance Theory (ART) [14, 15] and fuzzy c-mean clustering for medical classification and diagnosis with missing features was described.

Neural networks have also been applied to fire safety engineering [9–12]. Owing to the ability of neural networks in capturing non-linear system behaviors and the fast computational speed in making predictions, neural network models have been an alternative for simulating fire behaviors and learning fire dynamics. The results have confirmed the applicability of neural networks which have shown superior performances as compared with those from traditional models.

The work presented in this paper is focused on a new ART-based neural network model and its application to fire safety engineering. In [13], we proposed a Generalized Adaptive Resonance Theory model. GART is a hybrid ART and the Generalized Regression Neural Network (GRNN) [16] model that is compliant with Bayesian theorem. In this paper, the capabilities of GART are further enhanced. The objectives of this paper are two-fold. First, an enhanced GART network, known as O-EGART-PR, is proposed. O-EGART-PR is equipped with a Laplacian likelihood function, a new vigilance function, a match-tracking mechanism, an ordering algorithm for tackling the problem associated with the presentation sequence of training samples, and an ability to extract IF-THEN rules from the network weights. The second objective is to demonstrate the effectiveness of O-EGART-PR in handling problems in fire safety engineering. Specifically, O-EGART-PR is applied to predict the occurrence of flashover in compartment fire based on the fire size and the geometry of the compartment.

The organization of this paper is as follows. Section II describes the O-EGART-PR model. The experimental studies and comparison of results with other methods based on the prediction of flashover in compartment fires are given in Section III. A summary of this paper and suggestions for further work are presented in Section IV.

## II. THE O-EGART-PR MODEL

Figure 1 shows the architecture of the proposed O-EGART-PR model. The detail dynamics are given as follow.

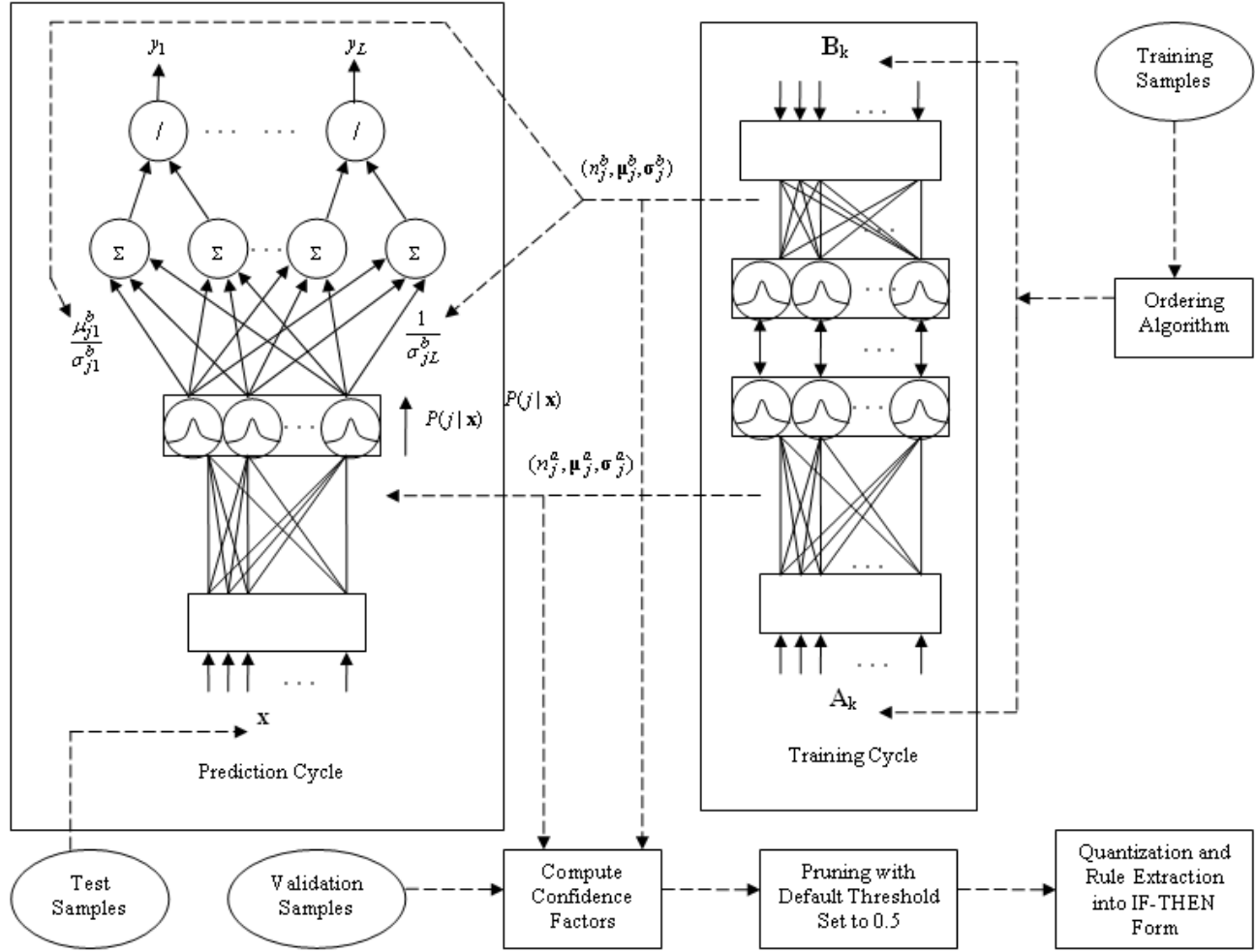


Figure 1 The architecture of the proposed O-EGART-PR model

### Sequence of Training Samples

The ordering algorithm originally proposed in [17] for ARTMAP-based network is used to determine the presentation sequence of training samples. The algorithm requires a predefined number of cluster centers,  $\omega$ . There are three stages in the ordering algorithm, as follows.

Stage 1 – Identify the first cluster center (the first training sample in the sequence): For each of the  $M$ -dimension input vector,  $\mathbf{A}_k$ , find the respective complement-coded [15] vector,  $\mathbf{I}_k \in \mathbf{R}^{2M}$  using

$$\begin{aligned} \mathbf{I}_k &= (\mathbf{A}_k, \mathbf{A}_k^c) = (\mathbf{A}_k, 1 - \mathbf{A}_k) \\ &\equiv (A_{k,1}, \dots, A_{k,M}, 1 - A_{k,1}, \dots, 1 - A_{k,M}) \\ &\equiv (A_{k,1}, \dots, A_{k,M}, A_{k,M+1}, \dots, A_{k,2M}). \end{aligned} \quad (1)$$

The  $K^{\text{th}}$  input vector that has the largest value as in Equation (2) is selected as the first sample in the presentation sequence.

$$K = \arg \max_k \left( \sum_{i=1}^M |A_{k,M+i} - A_{k,i}| \right) \quad (2)$$

Stage 2 – Identify the  $2^{\text{nd}}$  to  $\omega^{\text{th}}$  cluster centers, i.e., the  $2^{\text{nd}}$  to  $\omega^{\text{th}}$  training samples in the presentation sequence. The Euclidean distances of all remaining input vectors and the existing cluster center(s) are computed. For each input vector, the distance of the input vector and the nearest cluster center (i.e., the minimum distance) is then identified. The input vector that has the maximum value of these distances is selected as the next cluster center. The same procedure is repeated for the rest of the cluster centers.

Stage 3 – The sequence of the remaining input vectors is determined based on the minimum Euclidean distance from the input vectors to the cluster centers.

### Training Samples

Assume that the training samples presented to ART-a and ART-b are  $\{(\mathbf{A}_1, \mathbf{B}_1), (\mathbf{A}_2, \mathbf{B}_2), \dots, (\mathbf{A}_k, \mathbf{B}_k)\}$ , where  $\mathbf{A}_k \in \mathbf{R}^M$  and  $\mathbf{B}_k \in \mathbf{R}^1$  are the input vector and kernel label of the  $k^{\text{th}}$  training sample, and  $M$  is the number of attributes (or features) of the input vector, respectively. Note that in the following discussion, the equations and variables are based on ART-a with input sample  $\mathbf{A}_k$ . The equations and variables of ART-b with kernel label  $\mathbf{B}_k$  are the same but with subscript/superscript of “b” instead of “a”.

### Competition

The input sample is presented to ART-a with its kernel label to ART-b for computing the *choice and match* functions. The choice and match functions are defined based on the Bayesian theorem. The posterior probability of category- $j$  in ART-a to input sample  $\mathbf{A}_k$  is

$$P(j | \mathbf{A}_k) = \frac{P(\mathbf{A}_k | j)P(j)}{P(\mathbf{A}_k)}. \quad (3)$$

The prior probability is

$$P(j) = \frac{n_j^a}{\sum_{i=1}^N n_i^a} \quad (4)$$

and

$$P(\mathbf{A}_k) = \sum_{i=1}^N P(\mathbf{A}_k | i)P(i), \quad (5)$$

where  $P(\mathbf{A}_k | j)$  is a Laplacian likelihood function used to measure the similarity between  $\mathbf{A}_k$  and category- $j$ , and is defined as

$$P(\mathbf{A}_k | j) = \frac{1}{2^M \prod_{i=1}^M \sigma_{ji}^a} \exp\left(-\sum_{i=1}^M \frac{1}{\sigma_{ji}^a} |\mu_{ji}^a - A_{ki}|\right), \quad (6)$$

where  $\mu_j^a$ ,  $\sigma_j^a$  and  $n_j^a$  are the center, standard deviation, count of category- $j$  of ART-a, respectively. Note that while Gaussian ARTMAP [18] uses a standard quadratic loss function, O-EGART-PR employs a Laplacian loss function. This is because the quadratic loss function can be affected by outliers, especially in the presence of noisy data, which may lead to inaccurate recognition [19]. In order to find the “first round winner” of the competition, two measures are computed: the choice function of ART-a as defined in (1) and the match function as follows

$$V(\mathbf{A}_k, j) = \left(2^M \prod_{i=1}^M \sigma_{ji}^a\right) P(\mathbf{A}_k | j). \quad (7)$$

The first round winner of ART-a, which is denoted as  $J$ , is selected based on the highest value of the choice function, and its match function must be larger or equal to the vigilance parameter, i.e.,

$$J = \arg \max_j (P(j | \mathbf{A}_k)), \quad (8)$$

$$V(\mathbf{A}_k, J) > \rho_a. \quad (9)$$

where  $\rho_a \in (0,1)$  is a user-defined vigilance parameter of ART-a.

### Match Tracking

Once the first round winner of ART-a is found, verification is needed before it can be declared as the “final winner”. The vigilance test of ART-b is conducted, i.e.,

$$V(\mathbf{B}_k, J) \geq \rho_b, \quad (10)$$

where  $\rho_b \in (0,1)$  is a user-defined vigilance parameter of ART-b. If Eq. (10) is satisfied, category- $J$  is declared as the final *winner*, and is subject to learning. Otherwise, a match tracking mechanism is triggered to search for a better candidate to be the new first round winner from the existing categories in ART-a. During match tracking, the first round winner that fails to satisfy Eq. (10) is temporarily deactivated (with its choice function  $P(j | \mathbf{A}_k) = -1$ ), and  $\rho_a$  is temporarily increased to  $\rho_a = V(\mathbf{A}_k, J)$ .

### Learning

Learning involves the adjustment of the center, standard deviation, and counts of the final winner using the following equations.

$$n_J^a \leftarrow n_J^a + 1, \quad (11)$$

$$\mu_J^a \leftarrow \left(1 - \frac{1}{n_J^a}\right) \mu_J^a + \frac{\mathbf{A}_k}{n_J^a}, \quad (12)$$

$$\sigma_J^a \leftarrow \left(1 - \frac{1}{n_J^a}\right) \sigma_J^a + \frac{1}{n_J^a} |\mu_J^a - \mathbf{A}_k|, \quad (13)$$

### Adding a New Category

When none of the existing categories is able to fulfill Eq. (9) and pass the vigilance test as in Eq. (10), a new category is created to learn the new sample, i.e.,

$$N \leftarrow N + 1, \quad (14)$$

$$\mu_N^a = \mathbf{A}_k, \quad (15)$$

$$\sigma_N^a = \gamma_a, \quad (16)$$

$$n_N^a = 1, \quad (17)$$

where  $\gamma_a$  is a user-defined initial standard deviation value.

During the prediction cycle, an unlabeled sample,  $\mathbf{x}$ , is presented to ART-a, and a prediction is obtained based on a distributed posterior probability estimate using the GRNN procedure, as follows

$$f(\mathbf{x}) = \frac{\sum_{j=1}^N \frac{\mu_j^b}{\sigma_j^b} P(j | \mathbf{x})}{\sum_{j=1}^N \frac{1}{\sigma_j^b} P(j | \mathbf{x})}, \quad (18)$$

### Network Pruning

After the training phase is completed, the number of categories is reduced by a pruning procedure. Pruning aims to reduce the complexity of the network by removing those categories that have insignificant contributions to the overall

network output [20]. Another benefit of pruning is to facilitate the extraction of a compact rule set [21]. The network pruning procedure follows that in [22, 23]. A validation set is used to generate the confidence factors of each category. Consider an O-EGART-PR network that has been trained and validated with the training and validation sets. The confidence factor of category- $j$  is defined as

$$C_j = (U_j + R_j)/2, \quad (19)$$

where  $U_j$  and  $R_j$  are usage and accuracy of category- $j$ , respectively.

Unlike other neural network models that employ the “winner-take-all” strategy, O-EGART-PR works with probabilistic information. Hence, the definitions of  $U_j$  and  $R_j$  as in [22, 23] are not suitable for O-EGART-PR. Thus,  $U_j$  is defined as the total posterior probability in the validation set predicted by category- $j$ , divided by the maximum of the total posterior probability predicted by any category with the same label as category- $j$ , i.e.,

$$U_j = \frac{\sum_{k=1}^K P(j | \mathbf{x}_k)}{\max_{\text{for } i=1,2,\dots,N} \left\{ \sum_{k=1}^K P(i | \mathbf{x}_k), \text{ if } \mu_i^b = \mu_j^b \right\}}, \quad (20)$$

where  $K$  is the number of validation samples and  $N$  is the number of categories,  $\mu_i^b$  and  $\mu_j^b$  are the class labels for category- $i$  and category- $j$ , respectively. On the other hand,  $R_j$  is defined as the total posterior probability in the validation set *correctly* predicted by category- $j$ , divided by the maximum of the total posterior probability *correctly* predicted by any category with the same label as category- $j$ , i.e.,

$$R_j = \frac{\sum_{k=1}^K [P(j | \mathbf{x}_k), \text{ if } \mu_j^b = \Omega_k]}{\max_{\text{for } i=1,2,\dots,N} \left\{ \sum_{k=1}^K [P(i | \mathbf{x}_k), \text{ if } \mu_j^b = \Omega_k], \text{ if } \mu_i^b = \mu_j^b \right\}} \quad (21)$$

where  $\Omega_k$  is the class label of the  $k^{\text{th}}$  validation sample. The confidence factor is an index to indicate the category quality. A category with a high confidence factor implies a category that is frequently used and that has a high classification accuracy rate. Any category with a confidence factor smaller than a user defined threshold,  $\tau$ , is then pruned. For simplicity, the threshold value is set to  $\tau=0.1$  in this work.

#### Rule Extraction

After pruning the categories with low confidence factors, the remaining categories are used to extract fuzzy rules in the format of IF-THEN. The rule extraction method follows that in [22, 23]. A quantization approach is first applied. For the following discussion, the quantization level is set to 5, i.e.,  $Q=5$ . With five equalizations levels, the input attributes are described with five linguistic terms, i.e., “very low”, “low”, “medium”, “high” and “very high” (denoted as 1, 2, 3, 4, and 5). The quantized level of an input feature is computed as

$$V_q = \frac{q-1}{Q-1}, \quad (22)$$

where  $q = 1, 2, 3, 4, 5$ . In [22, 23], a category is represented by a hyperbox that has a lower bound and an upper bound, and the bounds are rounded to the nearest  $V_q$  level before being converted to linguistic terms. A category in O-EGART-PR is represented by a Laplacian likelihood function. Hence, the method in [22, 23] cannot be applied directly because there are no clear definitions of the lower and upper bounds in a Laplacian likelihood function.

A new method to find the lower and upper bounds of a category represented by a Laplacian likelihood function is needed. It is proposed to define the bound from the centre to left and right sides covering a certain percentage of the total likelihood. For the sake of simplicity, this region is empirically set to 50%, as shown in Fig. 2, in order to obtain a reasonable generalization. Hence, the lower and upper bounds ( $\theta_1$  and  $\theta_2$ ) are:

$$\theta_1 = \mu + \sigma \ln(0.5), \quad (23)$$

$$\theta_2 = \mu - \sigma \ln(0.5). \quad (24)$$

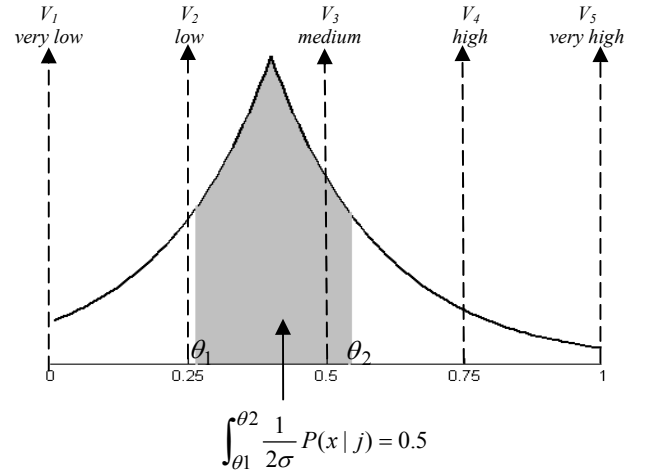


Fig. 2. The method to identify the bounds of a Laplacian likelihood function.

Once  $\theta_1$  and  $\theta_2$  are found, they are rounded to the nearest  $V_q$  and then converted into a linguistic term. For example, let two categories of O-EGART-PR with centers, standard deviations, target classes, and confidence factors, respectively be

$$\mu^a = \begin{pmatrix} 0.4 & 0.7 \\ 0 & 0.6 \end{pmatrix}, \sigma^a = \begin{pmatrix} 0.2 & 0.1 \\ 0.1 & 0.3 \end{pmatrix}, C = (1 \ 0.9)^T, \text{ the rules}$$

extracted from the categories are as follows.

*Rule 1:* If  $x_1$  is from low to medium and  $x_2$  is high, Then Class is  $-1$  with confidence factor  $C_1 = 1$ .

*Rule 2:* If  $x_1$  is very low and  $x_2$  is from medium to high, Then Class is  $+1$  with confidence factor  $C_2 = 0.9$ .

Note that, for simplicity, three O-EGART-PR parameters were set to their “default” values in this work, i.e.,  $\rho_a=0$ ,  $\rho_b=1$ , and  $\gamma_b = 1$ . Tuning of  $\gamma_a$  and  $\omega$  is conducted by trial-and-error in order to produce the best results.

### III. PREDICTING THE OCCURRENCE OF FLASHOVER IN COMPARTMENT FIRE

In this section, the applicability of O-EGART-PR is evaluated using a problem in fire-safety engineering. In general, it is expensive to obtain real fire samples from full-scale experiments. Hence computational fire simulations created by computer software are usually employed to generate data samples. These data samples are used for the training and testing of O-EGART-PR. In this study, the fire compartment is assumed to be rectangular in shape with an open door as illustrated in Fig. 3. The interaction between fire and the environmental parameters has been proposed in [24]. In this model, the temperature of the upper hot gas layer is a function of the room geometry, including the dimensions of the opening, the properties of the gas, the wall conduction characteristics, and the heat release rate. The criteria for flashover as defined by [25] were used in this study, and were input into the computer package FASTLite [26] to estimate the occurrence of flashover. The engine for this computer package is FAST [27]. In the study, a binary classification problem was studied. The following inputs were randomly generated, and used to train O-EGART-PR:

- room length (varied randomly from 2 to 10 m);
- room width (varied randomly from 2 to 10 m);
- room height (varied randomly from 2 to 10 m), and
- maximum heat release rate (varied randomly from 10 to 6000 kW).

The output was either flashover or non-flashover. Fast growth t-square ( $t^2$ ) fire [28] was also assumed through out the simulation. The growth of  $t^2$  fire is described by

$$\dot{Q} = \frac{dQ}{dt} = \alpha(t - t_0)^2 \quad (25)$$

where  $\dot{Q}$  is the heat released rate ( $kW$ ),  $\alpha$  is the growth constant ( $0.0469 \text{ kW s}^{-1}$ ),  $t_0$  is the initial time (s) and  $t$  is the time (s). The ceiling and walls in the fire simulation were assumed to be made of 16-mm gypsum, and the floor was assumed to be 12.7-mm plywood. For different combinations of room dimensions and maximum heat release rates, the occurrences of flashover as determined by FASTLite [26] were recorded for network training and testing. A total of 375 samples were generated, of which 190 were flashover samples and 185 were non-flashover samples.

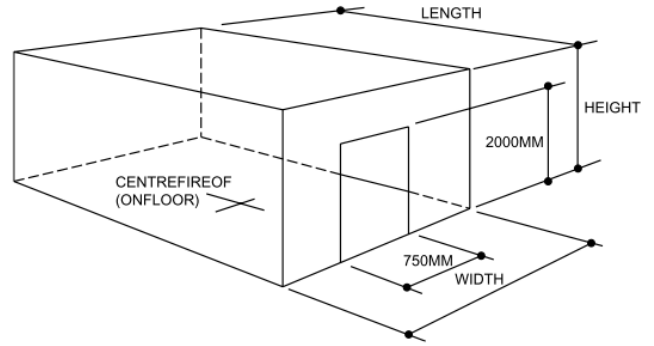


Fig. 3. Fire compartment for modeling the occurrence of flashover.

In accordance with [29], in each run, 250 samples were randomly selected for training and the rest were used for testing. A total of 50 runs were conducted, and the average results based on bootstrap means [30] (with 1,000 resamplings) are reported, as in Table 1. For comparison purposes, the results of GRNNFA, FAM and PEMap (extracted from [29]) are also included in Table 1. It is clear that O-EGART-PR outperforms other models. Out of the 50 runs, the one with the best prediction rate was used for rule extraction purposes. Note that in this case, the test set was used as the validation set for rule extraction. All the extracted rules from O-EGART-PR are given in Table 2.

TABLE 1  
PERFORMANCE COMPARISON ON THE PREDICTION OF FLASHOVER IN A COMPARTMENT FIRE

Model	Accuracy (%)
O-EGART-PR	94.48
GRNNFA [29]	92.60
FAM [29]	91.20
PEMap [29]	91.80

TABLE 2  
RULES EXTRACTED FROM O-EGART-PR WITH QUANTIZATION LEVEL  $Q = 5$  FOR THE FLASHOVER PREDICTION PROBLEM.

	If Attribute is				Then Class is	Confidence Factor
	i	ii	iii	iv		
Rule 1	2-4	2-4	2-4	1-2	-1	1
Rule 2	2-4	2-3	2-4	3-4	+1	1
Rule 3	2-3	3-4	1-2	2	+1	0.1335

Attribute i-iv: room length, room width, room height, maximum heat release rate. Quantization Level: 1 (very low); 2 (low); 3 (medium); 4 (high); 5 (very high). Classes: -1 (no flashover); +1 (flashover).

The rules given in Table 2 can be interpreted using the IF-THEN format. As an example, Rule 3 can be interpreted as:

*Rule 3: If room length is from low to medium and room width is from medium to high and room height is from very low to low and maximum heat transfer rate is low then the*

occurrence of flashover is true with confidence factor of 0.1335.

#### IV. CONCLUSIONS AND FUTHER WORKS

This paper has proposed a new GART model known as O-EGART-PR that possesses the ability of pattern classification as well as rule extraction. The performances of O-EGART-PR have been evaluated with a fire-safety engineering problem, i.e., predicting the occurrences of flashover in a compartment fire. The results have shown that O-EGART-PR is able not only to produce high accuracy rates as compared with other models but also to yield IF-THEN rules for explaining its predictions.

Although the results obtained from the experiments are encouraging, more studies using data sets from various application domains are needed. In addition, an ensemble method, which is able to increase the accuracy rates by voting, can also be investigated. Besides, a hybrid system using the genetic algorithm to optimize the performance of O-EGART-PR model can also be pursued as further work.

#### REFERENCES

- [1] P. Devijver, P. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice-Hall, NJ, 1982.
- [2] R.O. Duda, P.E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
- [3] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, 1972.
- [4] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995.
- [5] N. Belacel, P. Vincke, J.M. Scheiff, M.R. Boulassel, "Acute leukemia diagnosis aid using multi criteria fuzzy assignment methodology", *Journal of Computer methods and programs in biomedicine*, vol. 64, no. 2, pp. 145–151, 2001.
- [6] D. West, V. West, "Model selection for a medical diagnosis decision support system: a breast cancer detection case", *Journal of Artificial Intelligence in Medicine*, vol. 20, no. 3, pp.183–204, 2000.
- [7] P. Meesad, G.G. Yen, "Combined numerical and linguistic knowledge representation and its application to medical diagnosis", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 33, no. 2, pp. 206–222, 2003.
- [8] C. P. Lim, J. H. Leong, and M. K. Kuan, "A hybrid neural network system for pattern classification tasks with missing features", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 648–653, 2005.
- [9] Y. Okayama, "A primitive study of a fire detection method controlled by artificial neural net", *Fire Saf J*, vol. 17, no. 6, pp. 535–53, 1991.
- [10] H. Ishii, T. Ono, Y. Yamauchi, S. Ohtani, "Fire detection system by multilayered neural network with delay circuit", *Fire safety science – proceedings of the fourth international symposium*, pp. 761–72, 1994.
- [11] J.A. Milke, T. J. Mcavoy, "Analysis of signature patterns for discriminating fire detection with multiple sensors", *Fire Technol*, vol. 31, no. 2, pp. 120–36, 1995.
- [12] G. Pfister, "Multisensor/multicriteria fire detection: a new trend rapidly becomes state of art", *Fire Technol*, vol. 32, no. 2, pp.115–39, 1997.
- [13] K. S. Yap, C. P. Lim, and I. Z. Abidin, "A hybrid ART-GRNN online learning neural network with a  $\epsilon$ -insensitive loss function," *IEEE Trans. on Neural Networks*, vol. 19, no. 9, pp. 1641–1646, 2008.
- [14] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Journal of Neural Networks*, vol. 4, no. 6, pp. 759–771, 1991.
- [15] G. A. Carpenter, S. Grossberg, N. Markuzon, and J. H. Reynolds, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. on Neural Networks*, vol. 3, no. 5, pp. 698–713, 1992.
- [16] D. F. Specht, "A general regression neural network," *IEEE Trans. on Neural Network*, vol. 2, no. 6, pp. 568–576, 1991.
- [17] I. Dagher., M. Georgiopoulos, G. L. Heileman and G. Bebi, "An ordering algorithm for pattern presentation in fuzzy ARTMAP that tends to improve generalization Performance," *IEEE Trans. on Neural Networks*, vol. 10, no. 4, pp. 768–778, 1999.
- [18] J. R. Williamson, "Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps," *Journal of Neural Networks*, vol. 9, no. 5, pp. 881–897, 1996.
- [19] W. Chu, S. S. Keerthi, and C. J. Ong, "Bayesian support vector regression using a unified loss function," *IEEE Trans. on Neural Networks*, vol. 15, no. 1, pp. 29–44, 2004.
- [20] G.B. Huang, P. Saratchandran, and N. Sundararajan, "An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks," *IEEE Trans. on Systems, Man and Cybernetic Part B*, vol. 34, no. 6, pp. 2284–2292, 2004.
- [21] G. A. Carpenter and A. H. Tan, "Rule extraction: From neural architecture to symbolic representation", *Journal of Connection Science*, vol. 7, no. 1, pp. 3–27, 1995.
- [22] A. Quteishat, and C. P. Lim, "A modified fuzzy min-max neural network with rule extraction and its application to fault detection and classification," *Journal of Applied Soft Computing*, vol. 8, no. 2, pp. 985-995, 2008.
- [23] S. C. Tan, and C. P. Lim, "Application of an adaptive neural network with symbolic rule extraction to fault detection and diagnosis in a power generation plant," *IEEE Trans. on Energy Conversion*, vol. 19, no. 2, pp. 369–377, 2004.
- [24] B. J. McCaffrey, J. G. Quintiere, M. F. Harkleroad, "Estimating room fire temperatures and the likelihood of flashover using fire test data correlations", *Fire Technol*, vol. 17, no. 2., pp. 98–119, 1981.
- [25] V. Babrauskas, R. D. Peacock, P. A. Reneke, "Defining flashover for fire hazard calculations: part II", *Fire Saf J*, vol. 38, no. 7, pp. 613–22, 2003.
- [26] R. W. Portier, R. D. Peacock, P. A. Reneke, "FASTLite: engineering tools for estimating fire growth and smoke transport", *Special publication 899, National Institute of Standards and Technology*, 1996.
- [27] W. W. Jones, R. D. Peacock, "Technical reference guide for FAST version 18", *National Institute of Standards and Technology*, 1989.
- [28] G. Heskestad, "Engineering relations for fire plumes", *Society of Fire Protection Engineers, Technology report*, pp. 82-88, 1982.
- [29] E. W. M. Lee, Y.Y. Lee, C.P. Lim, C.Y. Tang, "Application of a noisy data classification technique to determine the occurrence flashover in compartment fires", *Advance Engineering Informatics*, vol. 20, pp. 213–222, 2006.
- [30] B. Efron, "Nonparametric standard errors and confidence intervals," *Can.J. Stat.*, vol. 9, pp. 139–172, 1981.