Journal of Administrative and Economic Sciences • Vol 16, Issue 2 (2023) • p 85-98

Support Vector Machines and Fuzzy Nonlinear Regression for Intelligent Identification of Urban VANET Constraints

[1,2] Bechir Alaya [3*] Anis Omri [1] Fahad Alaieri

[1] Department of Management Information Systems and Production Management, College of Business and Economics, Qassim University, 6633, Buraidah, 51452 • kingdom of Saudi Arabia

[2] Commande Numérique des Procédés Industriels Laboratory (CONPRI), ENIG, Gabes University • TUNISIA

[3] Department of Business Administration, College of Business and Economics, Qassim University, 6633, Buraidah, 51452 • kingdom of Saudi Arabia

*Coorsponding author: a.omri@qu.edu.sa

(Received: September 18, 2022 • Revised: October 05, 2022 • Accepted: October 10, 2022)

Abstract:

Nowadays, video-on-demand (VoD) applications are becoming one of the tendencies driving vehicular network users. In this paper, considering the of ve-hicular network opens up to different types of communications in order to meet the needs of the wide variety of new applications envisaged within the framework of the Intelligent Transport System (ITS). In this work, we seek to establish a list of possibilistic concepts in order to efficiently identify the strict parameters of ur-ban VANET networks. To this end, we use linear optimization under constraints. We apply in parallel to this first proposition a minimization of a validated quadrat-ic criterion with the appearance of fuzzy least squares. To arrive at a quadratic resolution under constraints, different distances were managed and various con-straints were introduced in the optimization problem. We have shown that the da-ta independent criterion in urban VANETs can overcome the failure problem in terms of robustness. To assess the comparative effectiveness of our solutions, many experiments are carried out. The obtained results showed that the proposed identification scheme will allow an increase in the performance of Urban VANET networks with different load conditions.

Keywords: Vehicular network, SVM, UPSO, regression, optimization, parameters identification.

1. Introduction

For several years, research work on regressive models with fuzzy parameters (fuzzy regression) has multiplied, and it is now very difficult to establish an exhaustive list of all the strategies developed. The basic idea was to exploit possibilistic concepts to identify the fuzzy events of VANET networks by minimizing the dispersion of the coefficients. A linear optimization program under constraints then translates the strategy implemented to achieve this identification. The constraints are dictated by the realization of inclusion of the outputs observed in those predicted by the model (known as the possibility model).

In parallel with these works, the idea of identification of continuous behavior of the vehicle nodes in a VANET environment based on the minimization of a quadratic criterion materializes with the appearance of the fuzzy least squares [1]. Because of its great ability to adapt and integrate with the most precise identification methods, it seems quite natural to us to exploit the idea of the least-squares in our uncertain VANET network context. Indeed, the underlying guiding idea is to minimize the quadratic error of the output, expressed in terms of the distances between the fuzzy intervals (observed outputs and expected outputs). In this context, different distances were used as criteria to be optimized (diamond distance, Ming distance, ...) [2].

However, if this method often gives relevant results, thanks to its principle, it is often subject to the violation of the inclusion constraint. In this context, to remedy these problems, constraints have been introduced into the optimization problem to lead to a quadratic programming methodology under constraints [3]. Our job is to perform an experimental analysis of the differences that exist between conventional approaches in terms of the robustness of VANET systems. To do this, we have chosen to approach the problem practically by relying on simulations making it possible to highlight the sensitivity of each method at any modification of the observed data, or a change in the data learning area.

The rest of this paper is organized as follows; section 2 outlines a literature review on the fuzzy logic approach applied in the VANET environment. Section 3 describes our adaptation Fuzzy Support Vector Regression for VANET networks (FSVRNET). In section 4 we propose a new fuzzy identification approach based on Support Vector Fuzzy Regression (FSVRNET) and the Unified Particle Swarm Optimization (UPSO). Finally, an analysis of the experience study is presented in section 5. Section 6 concludes the paper.

2. Review of related work

Support Vector Machines (SVM) offer the possibility of performing a linear regression to no longer predict a class, but any function [4]. Assuming that the functions with values in IR, then N solutions have been provided, see in [5] a simple regression is performed on each of the dimensions. With functions of belief or membership functions, the normalization constraints require taking into consideration the different dimensions of the function to be predicted independently and jointly [6]. In [7], the authors have proposed a regression on triangular fuzzy numbers, it is thus placed in the hypothesis of multiple regression, but the essential difference with the proposed approach is manifested in the conditions on the belief functions and the membership functions.

To meet various network resource requirements, VANET networks then require efficient approaches for scalable resource allocation. In a VANET environment and to meet the resource requirements, the authors of [8] presented the H-FLGA approach (Hybrid Fuzzy Logic Approach which expresses five ways of optimizing network resources.

For efficient transmission of various data in a VANET environment, a self-adaptive geographic routing has been proposed in [9], with a hybrid reactive mechanism (SOGR-HR). The authors of [9] also proposed a self-adapting geographic routing, but with a geographic reactive mechanism (SOGR-GR).

The authors of [10] introduced "MARS", a system based on machine learning for the selection of traffic in urban VANET networks. A widely applied K-means machine learning algorithm is adopted by MARS to predict the better transmission capacity of selected paths. Another mechanism named FLGR based on fuzzy logic has been proposed in [11] for transmission control with a minimum delay of security messages. FLGR uses fuzzy logic and several parameters to select the best vehicle node that will act as the relay node.

One of the promising technologies in complex and dynamic contexts is fuzzy logic [12]. To improve the quality of data transmission (DTQ), the authors of [13] proposed a new data transmission strategy based on fuzzy logic and machine learning algorithms to optimize the selection of nodes relay. To solve the problems of the dynamics of the VANET, each RSU will estimate the movements of the vehicle nodes and their directions using the KNN algorithm and a machine learning system (MLS).

3. Fuzzy Support Vector Regression for VANET networks (FSVRNET)

In the systems modeling where the available information is uncertain as in the case of the VANET environment, we are dealing with a fuzzy approach to the system considered. A fuzzy function whose parameters represent fuzzy intervals expresses the system [14]. In the rest of this section, we integrate the main concepts of fuzzy logic in the regression model with SVMs in the VANET environment. In the support vector regression model, the parameters were set so that the desired results, the components of the weight vector, and the bias term are fuzzy numbers. To simplify the calculation, we will assume that the fuzzy parameters are symmetric triangular fuzzy numbers (NFTS) [15].

3.1 Quadratic function (QF) applied to VANET networks

First, the bias term and the weight vector components used in the regression function model are NFTS. Given the fuzzy weight vector $\mathrm{FW}_i = (\mathrm{m}_i, r_i)$, where m the midpoint, and r the radius, likewise the fuzzy bias term $Fb = (e, k)$, where, e midpoint, and k the radius. The model is given by:

$$W = \mathrm{FW}_1\mathrm{y}_1 + \mathrm{FW}_2\mathrm{y}_2 + \cdots + \mathrm{FW}_n y_n = \langle \mathrm{FW.\,y} \rangle + Fb$$

The following membership function defines the model:

$$\sigma_W(x) = 1 - \frac{|x - (\langle FW.y \rangle + e)}{\langle r.|y| + k \rangle}, \quad for\ all\ y \neq 0 \tag{1}$$

Such as $\sigma_W(x) = 0$, when $\langle \mathrm{r}.|\mathrm{y}| + \mathrm{k} \rangle \leq |x - (\langle \mathrm{FW.y} \rangle + Fb)|$

Then, we process the desired fuzzy output in the regression task of our VANET system. The observed data noted $\widetilde{W_i} = (x_i, \mathrm{b}_i)$, are also fuzzy values, where b_i the radius and x_i is the midpoint.

We then distinguish the membership function $\widetilde{W_i}$ as follows:

$$\sigma_{\widetilde{W_i}}(x) = 1 - \frac{|x - x_i|}{\mathrm{b}_i} \tag{2}$$

To formulate our fuzzy linear regression model in our VANET environment, the following points were retained:

1) The data can be represented by a fuzzy linear model $\mathrm{W}_i^* = \langle \mathrm{FW}^*.\mathrm{y}_i \rangle + \mathrm{Fb}^*$, given y_i, W_i^* can be obtained as follows:

$$\sigma_{W_i^*}(x) = 1 - \frac{|x - (\langle FW.y_i \rangle + e)}{\langle r.|y_i| + k \rangle} \tag{3}$$

2) The adjustment degree of the estimated fuzzy linear model. $\mathbf{W}_i^* = \langle \mathbf{FW}^*.\mathbf{y}_i \rangle + \mathbf{Fb}^*$ for the observed data $\widetilde{\mathrm{W}_i} = (\mathrm{x}_i, \mathrm{b}_i)$, is measured by the index $\overline{\mathrm{g}_i}$, which maximizes g, subject to $\widetilde{\mathrm{W}_i}^{\mathrm{g}} \subset \mathrm{W}_i^{*\mathrm{g}}$, with:

$$\widetilde{W_i}^g = \{x | \sigma_{\widetilde{W_i}}(x) \geq g\}, W_i^{*g} = \{x | \sigma_{W_i^*}(x) \geq g\} \tag{4}$$

Which are sets by $g - cut$. The optimum adjustment degree for all data is defined by, $min_j |\ \overline{g_j}|$.

3) The imprecision of the linear fuzzy model is defined by: $\frac{1}{2}\|\mathrm{r}\|^2 + \mathrm{k}$.
4) The control term of the fuzzy linear model capacity is defined by: $\|\chi\|^2$.

Solving the regression problem gives the fuzzy weight vector $\mathrm{FW} = (\chi, r)$ and the fuzzy bias term $\mathrm{Fb} = (e, k)$, such as the adjustment degree between the estimated output $W_i^{*\mathrm{g}}$ and the output desired $\widetilde{W_i}^g$, is a given constant h for all i, where h selected by the user, as the adjustment degree of the fuzzy linear model. The $\overline{g_i}$ value can be obtained from:

$$\overline{g_i} = 1 - \frac{|x_i - (\langle FW.x_i \rangle + e)}{\langle r.|x_i| + k \rangle - b_i} \tag{5}$$

Which is the same as that obtained by Škrabánek [16]. Our regression problem is, therefore:

$$\min_{\chi, r, e, k, \eta_i} \Psi = \frac{1}{2}\|\chi\|^2 + K\left(\frac{1}{2}\|\chi\|^2 + k\right) + \mathcal{R}\sum_{i=1}^{N} \eta_i \tag{6}$$

$$Subject\ to: \overline{g_i} \geq h \qquad \text{with } i = 1, \ldots, N$$

Such that,, $\|\chi\|^2$ characterizes the model complexity, and $\frac{1}{2}\|\chi\|^2 + k$ characterizes the imprecision model.

More imprecision in the result of the regression generally implies more imprecision in the linear fuzzy regressive model. K represents the difference parameter chosen by the decision-maker. The value of h determines the lower bound of the establishment of the linear fuzzy model, and $\overline{g_i}$ is the adjustment degree of the estimated linear fuzzy model $W_i^* = \langle \mathrm{FW}^*. y_i \rangle + \mathrm{Fb}^*$ for the fuzzy data of the desired output $\widetilde{W_i} = (x_i, \mathrm{b}_i). \{\eta_i\}_{i=1,2,\dots,N}$ are sets of variables that measure the excess quantity of the variation in stresses for each point, where $\mathcal{R}$ is a fixed penalty parameter chosen by the user, a greater value of $\mathcal{R}$ corresponds to assigning a higher penalty to errors [16].

More precisely, according to (5), our problem consists in finding the fuzzy weight vector $\mathrm{FW}^* = (\chi, r)$ and the fuzzy bias term $\mathrm{Fb}^* = (e, k)$, which is the solution of the problem of following quadratic function:

$$\min_{\chi,r,e,k,\eta_i} \Psi = \frac{1}{2}\|\chi\|^2 + K\left(\frac{1}{2}\|r\|^2 + k\right) + \mathcal{R}\sum_{i=1}^{N}(\eta_{1i} + \eta_{2i}) \qquad (7)$$

The difference between the support vector fuzzy regression technique and the simple regression model with SVM is that the first technique looks for a linear fuzzy function with fuzzy parameters that have at least the degree of fit h of the desired fuzzy targets of all data. On the other hand, the second model searches for a linear function which has at most ϵ of variation compared to the targets obtained x_i for all the data.

3.2 Adaptation of a fuzzy nonlinear regression model to VANET environment

There are only a few papers on fuzzy nonlinear regression [17][18]. This model type generally assumes the underlying model and deal with the estimation procedures of some particular models such as linear, polynomial, and exponential. However, we believe that these methods are dependent on an impractical model. We will use the idea of SVM for simple nonlinear regression [19] in our VANET network system. The basic idea is to change the data space. The nonlinear transformation of the data in a urban VANET environment can allow a linear representation of the observed data y_i by $\Omega := \mathcal{R}^n \rightarrow \mathcal{F}$. An essential property of the proposed algorithm is that it would depend only on the data by the interior products in $\mathcal{F}$, on the functions of the form $\langle \Omega(y_i). \Omega(y_j) \rangle$ and $\langle \Omega(|y_i|). \Omega(|y_j|) \rangle$.

The form of the function Ω does not need to be known because it is implicitly defined by the choice of the kernel function, $\mathcal{K}(a, b) = \langle \Omega(a). \Omega(b) \rangle$. Therefore, it suffices to know that $\mathcal{K}(y_i, y_j) = \langle \Omega(y_i). \Omega(y_j) \rangle$ *and* $\mathcal{K}(|y_i|, |y_j|) = \langle \Omega(|y_i|). \Omega(|y_j|) \rangle$instead of explicitly defining $\Omega(.)$.

Thus, by substitution $\langle y_i. y_j \rangle$ and $\langle |y_i|. |y_j| \rangle$ with $\mathcal{K}(y_i, y_j)$ and $\mathcal{K}(|y_i|, |y_j|)$ respectively, we obtain the dual quadratic optimization problem given by the following function. Here, it should be noted that the constraints are not modified [20].

$$\sum_{i=1}^{N}(\beta_{1i}-\beta_{2i})=0,\sum_{i=1}^{N}(\beta_{1i}+\beta_{2i})\leq\frac{\mathcal{K}}{1-h}\qquad\beta_{1i},\beta_{2i}\in[0,\mathcal{R}]\tag{8}$$

We can arrive at various types of learning machines, based on kernel functions and arbitrary nonlinear regression functions in the input space of our VANET system. The fuzzy nonlinear regression function is defined by the following membership function:

$$\sigma_{W^*}(x)=1-\frac{\left|x-\sum_{j=1}^{N}(\beta_{1j}-\beta_{2j})\,\mathcal{K}\left(y,y_j\right)+e\right|}{\frac{(1-h)}{\mathcal{K}}\sum_{j=1}^{N}(\beta_{1j}+\beta_{2j})\,\mathcal{K}\left(|y|,|y_j|\right)+e'}\tag{9}$$

With a projection of the data in a space of important dimensions, the kernel functions offer alternative solutions to increase the computational capacity of the model. In its dual form, the fuzzy linear regression model can perform these measurements perfectly and the adjustable variables do not depend on the number of attributes used [21]. Learning algorithms can generally be detached from the constraints of the application domain. It must be programmed for the design of a suitable kernel function.

3.3 Application of FSVRNET to VANET environment

To evaluate the performance of our FSVRNET model to the VANET environment, we assume the data drawn from Tanaka and Lee [22] to apply it to VANET. Note here that the representation of intervals by their midpoint, radius has been adopted. The data set presented therefore comprises M = 8 samples. We assume the fuzzy linear regression model as follows: $W(x)=FW\ x+Fb$

We have set the parameters $h=0.5$, $\mathcal{R}=15$, and $\mathcal{K}=25$. We obtain the following regressive model:

We apply the method of nonlinear fuzzy regression by SVM to analyze the data which seems appropriate for this model. For the estimation of the fuzzy nonlinear model, we use in this case a polynomial kernel of degree three.

For nonlinear fuzzy SVM regression (FSVR), we do not get good results in the case of an RBF (Radial Basis Function) kernel. So we use in this case almost the same technique, but with unsymmetrical triangular fuzzy numbers.

4. Fuzzy identification approach based on FSVRNET and Unified Particle Swarm Optimization (UPSO)

In this section, we propose a new fuzzy identification approach based on Support Vector Fuzzy Regression (FSVRNET) and the Unified Particle Swarm Optimization (UPSO) [23] algorithm.

First, the FSVRNET model is chosen to facilitate the identification model in a VANET environment. Then, to optimize the hyper-parameters essential for the FSVRNET model, a global UPSO optimizer is implemented. Applying the proposed model to different data provided optimal results. First, we will present the techniques used in our new approach to explain the different stages of the latter.

In this section, we briefly discuss Support Vector Fuzzy Regression (FSVR) by Hong et al [24], which uses unsymmetrical triangular fuzzy numbers. To do this we need some preliminaries. Let $\mathcal{A} = (m, r_L, r_R)$ an asymmetric triangular fuzzy number. Knowing that, r_L, r_R the left and right radius, respectively, and m is the modal value of $\mathcal{A}$. We use the distance from Diamond d for unsymmetrical triangular fuzzy numbers.

4.1 Linear fuzzy regression model

Given a set of data pairs (x_i, Y_i), $i = 1, \dots, N$, where the outputs Y_i are symmetric triangular fuzzy numbers and x_i are net numbers.

Let the fuzzy weight vector $W(w_1, w_2, \dots, w_M)$, where, $w_j = (m_{w_j}, r_{L_{w_j}}, r_{R_{w_j}})$, $j = 1, \dots, M$, belongs to the set of unsymmetrical triangular fuzzy numbers, the fuzzy bias term $B = \left(m_B, r_{L_B}, r_{R_B}\right) \in asymmetric\ NFT$.

We consider the following fuzzy regression model:$f(x) = \langle W, x\rangle + B = w_1x_1 + \cdots + w_Mx_M + B$

We define$\|W\|^2 = \|m_w\|^2 - \left\|m_w - r_{L_w}\right\|^2 + \left\|m_w + r_{R_w}\right\|^2$,

with

$$m_w = \left(m_{w_1}, m_{w_2}, \dots, m_{w_M}\right), \qquad r_{L_w} = \left(r_{L_{w_1}}, \dots, r_{L_{w_M}}\right) \tag{10}$$

We use the Lagrange method to resolve this optimization problem, we derive the Lagrange equation concerning $m_W, r_{L_w}, r_{R_w}, m_B,\ r_{L_B},\ r_{R_B}, \xi_{li}, \xi^*{}_{li}$, we obtain:

$$m_W = \sum_{i=1}^{N}\left(\alpha_{1_i} - \alpha^*{}_{1i}\right)x_i, \qquad r_{L_w} = \sum_{i=1}^{N}\left[\left(\alpha_{1_i} - \alpha^*{}_{1i}\right) - \left(\alpha_{2_i} - \alpha^*{}_{2i}\right)\right]x_i$$

$$r_{R_w} = \sum_{i=1}^{N}\left[\left(\alpha_{3_i} - \alpha^*{}_{3i}\right) - \left(\alpha_{1_i} - \alpha^*{}_{1i}\right)\right]x_i \tag{11}$$

Such that: $\alpha_{l_i}, \alpha^*{}_{li}, l = 1, \dots, 3, and\ i = 1, \dots, N$ represents the not negative Lagrange multipliers. The two radius $r_{L_w}, and\ r_{R_w}$must be always defined as positives, $r_{L_w} \geq 0, and\ r_{R_w} \geq 0$.

We obtain the corresponding dual optimization problem as follows:

$$\underset{\alpha_{l_i}, \alpha^*{}_{li} \geq 0}{Maximize}\begin{cases} \displaystyle\sum_{i=1}^{N}\left(\alpha_{l_i} - \ \alpha^*{}_{li}\right) = 0\,, l = 1, \dots, 3 \\ \alpha_{l_i}, \alpha^*{}_{li} \in [0. C], l = 1,2,3, i = 1, \dots, N \end{cases} \tag{15}$$

We can write our fuzzy model as follows:

$$f(x) = \ \left(\langle m_W, x\rangle, \langle r_{L_w}, x\rangle, \langle r_{R_w}, x\rangle\right) + B \tag{16}$$

We need to find now, $m_B, r_{L_B},\ and\ r_{R_B}$.

With the use of Karush-Kuhn-Tucker (KKT) conditions, we can calculate m_B as follows:

$$\begin{cases} m_B = m_{Y_i} - \langle W, m_{x_i} \rangle - \varepsilon, \quad While \;\; \alpha_{1i} \in (0, C) \\ m_B = m_{Y_i} - \langle W, m_{x_i} \rangle + \varepsilon, \quad While \;\; \alpha^*{}_{1i} \in (0, C) \end{cases} \tag{18}$$

To find the value of r_{L_B}, $and \; r_{R_B}$, we need to solve the optimization problem given below $\underset{r_{L_B}, r_{R_B} \geq 0}{\text{Min}}$ of:

$$\begin{cases} \sum_{i=1}^{N} \left| m_{Y_i} - r_{L_{Y_i}} - \langle m_W - r_{L_W}, x_i \rangle - m_B + r_{L_B} \right|_\varepsilon \\ + \sum_{i=1}^{N} \left| m_{Y_i} + r_{R_{Y_i}} - \langle m_W + r_{R_W}, x_i \rangle - m_B + r_{R_B} \right|_\varepsilon \end{cases} \tag{19}$$

Such that the loss function $\epsilon - insensitive$ given below:

$$L_\epsilon(y) = \begin{cases} 0 \quad While \; |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon \quad otherwise \end{cases} \tag{20}$$

The combination of exploration and exploitation properties of two variants of PSO, local and global, [25], gave rise to the Unified Particle Swarm Optimization technique (UPSO). The UPSO technique is mainly based on the inertial version of PSO, although it can be defined directly for the constriction factor version. Let $L_i(t+1) \; and \; G_i(t+1)$ denote the speed adaptation of the i^{th} vehicle, for the local and global variant [26].

$$G_i(t+1) = V_i(t) + c_1 r_1 \big(P_i(t) - X_i(t)\big) + c_2 r_2 \left(P_g(t) - X_i(t)\right) \tag{21}$$

$$L_i(t+1) = V_i(t) + c_1 r'_1 \big(P_i(t) - X_i(t)\big) + c_2 r'_2 \left(P_{g_i}(t) - X_i(t)\right) \tag{22}$$

where t denotes the current iteration, g_i represents the index of the most suitable vehicles which are in the vicinity of X_i, while g represents the index of the most suitable vehicle of the swarm. By combining equations 21 and 22 into a single equation, we then obtain the main mechanism of UPSO.

$$V_i(t+1) = (1-u).\, G_i(t+1) + u.\, L_i(t+1) \tag{23}$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{24}$$

where the parameter u is the unification factor. This factor balances the influence of the global and local search directions in the final algorithm, the case of the global PSO standard is obtained by putting $u = 1$ in equation 23, $u = 0$ corresponds to the case of the local PSO standard. All the values of $u \in (0,1)$, corresponding to composing the variants of PSO which combine the exploration and exploitation characteristics of its local and global variant.

4.2 Hyper-parameters selection in VANET environment by the PSO and UPSO standards

The hyper-parameters of the FSVRNET model ($C \; and \; \sigma$) can be optimized by using the UPSO and PSO algorithms. Consequently, each vehicle represents a potential solution with the resolution of the problem of selecting hyper-parameters. The hyper-parameters are evaluated using a fitness function, which is determined concerning the considered optimization problem. The goal of FSVRNET's learning and testing process is to further develop the generalization performance of the fuzzy regression model [27] [28]. We can then define the fitness function as follows:

$$Fitness = \frac{1}{v}\sum_{i=1}^{v}\sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(m_{f(x_{ij})} - m_{Y_{ij}}\right)^2} \tag{25}$$

Where, v is the number of folds for the cross-validation of $v - folds$, and m the number of each validation subset, $m_{Y_{ij}}$ is the center of the fuzzy output value observed, $m_{f(x_{ij})}$ is the estimated fuzzy model midpoint of all learning data.

The objective is to select the optimal hyperparameters. Also to optimize the physical form of the vehicle, therefore, that it should be reserved during the optimization process.

5. Analysis of experience

To validate the feasibility and effectiveness of FSVRNET on the identification quality in a VANET environment, an experimental study is conducted through a multi-entry system. In this application, the dataset, composed of 100 samples. The size of the model is assumed to be known.

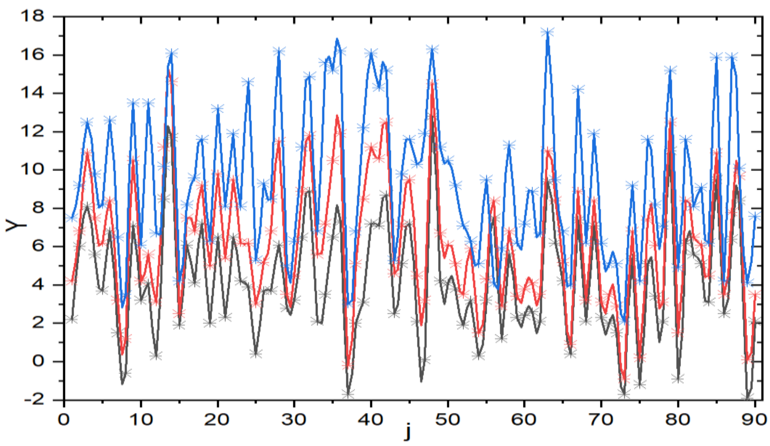

Figure 1: Multivariable model obtained by UPSO-FSVRNET (Model 1)

Generally, and as assumed in most identification methods, the model obtained is validated on a test set. If the identification data very small amount, it is possible to use statistical validation tools. First, a linear system of the form:

$$\hat{Y} = f(x) = B \oplus \sum_{i=1}^{N} W_i.(x_i - shift_i) \tag{26}$$

From the used data, it is clear that all the inputs are between 0 and 1. Therefore, first of all, the selected offsets are defined by $shift_i = 0, i = \{1, 2, 3, 5\}, and\ shift_4 = 1$.

We now apply our proposed PSO/UPSO-FSVRNET approach to this dataset. First, we consider a fuzzy model of the form 40 (Model 1) and we use the RBF (Radial Basis

Function) kernel to map the input data into the characteristic space. The model obtained (Model 1) is illustrated in Figures .2.

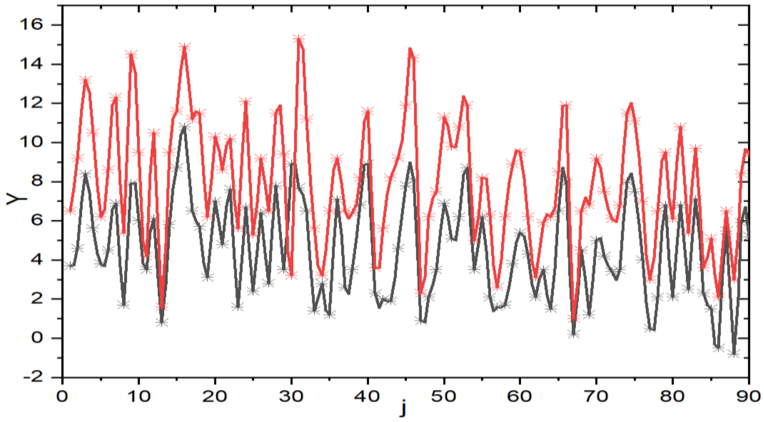

Figure 2: Model obtained by the linear approach (Model 1)

We pass to the task of validating this model obtained using the test data (cf. Annex-1). To do this, we must calculate the measure $MAPE$, and $MAPE_j$ (these equations are for triangular fuzzy models) as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{m_{f(x_i)} - m_{y_i}}{m_{y_i}}\right| \times 100 \qquad (27)$$

With m_{y_i} is the midpoint of the true fuzzy output, $m_{f(x_i)}$ is the midpoint of the estimated fuzzy regression model.

$$MAPE_j = \left|\frac{m_{f(x_j)} - m_{y_j}}{m_{y_j}}\right| \times 100 \qquad (28)$$

We have gathered the $MAPE_j$ measurement for each test data, and also the $MAPE$ error of all the test data, and for all identification data. For trapezoidal models, the $MAPE$ measure can be calculated from the follows equation:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{m_{S_{f(x_i)}} - m_{y_i}}{m_{y_i}}\right| \times 100 \qquad (29)$$

With $m_{S_{f(x_i)}}$ is the modal value of the support of the estimated fuzzy output (prediction), and m_{y_i} is the midpoint of the observed output. The $MAPE_j$ is calculated as follows:

$$MAPE_j = \left|\frac{m_{S_{f(x_i)}} - m_{y_j}}{m_{y_j}}\right| \times 100 \qquad (30)$$

Table 1: Test on validation data for model-1 by the linear technic.

j	y	$m_{f(x)}$	$MAPE_j$ (%)	$\hat{Y}$
1	1.3915	1.5300	9.9503	[0.2675, 1.8734]
5	1.1886	-0.0340	102.8584	[−1.9941, 0.4201]
26	1.7484	2.2275	27.4029	[1.6781, 2.9302]
32	2.3165	2.0436	11.7818	[0.5660, 2.9745]
34	5.1463	5.6391	9.5763	[3.8384, 5.7367]
41	2.2748	2.7977	22.9880	[1.4329, 3.4925]
67	1.4327	0.4190	70.7555	[-1.3283, 1.3375]
93	2.7117	2.4326	10.2926	[1.5477, 3.5363]
$MAPE$ (test data)			35.0011	
$MAPE$ (identification data)			14.3083	

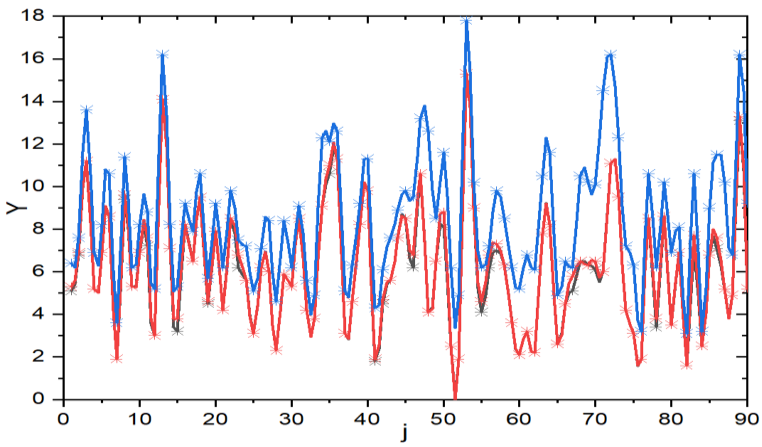

Figure 3: Multi-entry model obtained by UPSO-FSVRNET (Model 2).

Practically, the choice of structure is made by empirical considerations, which consist in trying several structures and accepting only the one that gives the best performance. Thus, it is possible to incorporate a technique for identifying the structure of the model before using our method.

The model to be identified is therefore of the form:

$$\hat{Y} = f(x) = B + W_1.(x_1.x_2) \oplus W_{2.}x_3 \oplus W_{3.}(x_3.x_4) \oplus W_{4.}(x_5)^2 \qquad (31)$$

The obtained models using the proposed method is shown in Figures 6.

6. Conclusion

In this paper, we have presented two types of methods applied in the identification of a fuzzy regressive model in a VANET environment. One is to minimize a linear criterion or quadratic depending on the observed data and/or the measured outputs, this is the conventional approach, while the other amounts to minimizing a criterion completely independent of the latter, it is the fuzzy support vector regression (FSVRNET). It should also be noted that if one wishes to obtain a possibility model, which respects the complete inclusion of the data in the predicted output, the minimization must be done under constraints in all cases, and as the complexity of a problem in linear optimization is less important, it makes more sense to use the proposed linear approach. Also, in this work, we proposed a new contribution of the identification of fuzzy regressive models based on the method of FSVRNET with unsymmetrical triangular fuzzy numbers and unified particle swarm optimization (UPSO), this new method was tested in two illustrative monovariable examples, then applied it to a noisy multi-input dataset.

Our evaluation was limited to a study of optimal identification performance and hyper-parameters for the two multi-input models studied. However, we did not evaluate the real-time performance of the entire system. Several perspectives are possible, such as the proposed approach being developed with the assumption that the proposed identification model uses all the necessary information from the different entities of the VANET network. The definition of the view of the network to be shared and how it will be shared, as well as the operating mode for exchanging data between the different actors of this system, can constitute a first architectural perspective. On the other hand, the definition of the East-West interface (defining the messages exchanged between the entities of the VANET network) taking into account the mobility constraints of the nodes represent a second architectural perspective to be explored. Taking a dynamic approach to smart identification holds great promise.

References

1. L. Sellami, B. Alaya, SAMNET: Self-adaptative multi-kernel clustering algorithm for urban VANETs, Vehicular Communications, Vol. 29, 2021.
2. F.G. Juan Carlos, Y.C. Yurilev, R.F. Heriberto, Distance measures for Interval Type-2 fuzzy numbers, Discrete Applied Mathematics, Vol. 197, pp- 93-102, 2015.
3. K. Kandali, L. Bennis, H. Bennis, A New Hybrid Routing Protocol Using a Modified K-Means Clustering Algorithm and Continuous Hopfield Network for VANET. IEEE Access. Vol. 9, pp. 47169-47183, 2021.
4. J. Cervantes, F.G. Lamont, L. Rodríguez, A. Lopez-Chau, Asdrubal. A comprehensive survey on support vector machine classification: Applications, challenges, and trends. Neurocomputing, Vol. 408, pp. 189-215, September 2020.
5. M. Habshah, M. Jama, A Modified Robust Support Vector Regression Approach for Data Containing High Leverage Points and Outliers in the Y-direction, Mathematics and Statistics, Vol. 8, No. 5, pp. 493 - 505, 2020.
6. M.S Ahmad, S.M. Adnan, S. Zaidi, Pr. Bhargava, A novel support vector regression (SVR) model for the prediction of splice strength of the unconfined beam specimens, Construction and Building Materials, Vol. 248, 2020.

7. Z. Xue, R. Zhang, C. Qin, X. Zeng, An adaptive twin support vector regression machine based on rough and fuzzy set theories, Neural computing & applications, Vol. 32, No. 9, pp. 4709-4732, 2020.
8. A. A. Khan, M. Abolhasan, W. Ni, J. Lipman, and A. Jamalipour, "A Hybrid-Fuzzy Logic Guided Genetic Algorithm (H-FLGA) Approach for Resource Optimization in 5G VANETs," in IEEE Transactions on Vehicular Technology, vol. 68, no. 7, pp. 6964-6974, July 2019.
9. X. Xiang, X. Wang, Z. Zhou, "Self-Adaptive On-Demand Geographic Routing for Mobile Ad Hoc Networks," IEEE Transactions on Mobile Computing, vol. 11 no. 9, pp. 1572-1586, DOI: 10.1109/TMC.2011.177, 2012.
10. W. K. Lai, M. Lin, Y. Yang, "A Machine Learning System for Routing Decision-Making in Urban Vehicular Ad Hoc Networks," International Journal of Distributed Sensor Networks, vol. 11 no. 3, 2015.
11. S. Agrawal, R. S. Raw, N. Tyagi, A. K. Misra, "Fuzzy logic based greedy routing (flgr) in multi-hop vehicular Adhoc networks," Indian Journal of Science and Technology, vol. 8 no. 30, 2015.
12. I. Tal, G. Muntean, "Towards Reasoning Vehicles," ACM Computing Surveys, vol. 50 no. 6, 2018.
13. Y. Zhou, H. Li, C. Shi, N. Lu, N. Cheng, A Fuzzy-Rule Based Data Delivery Scheme in VANETs with Intelligent Speed Prediction and Relay Selection. Wireless Communications and Mobile Computing, pp. 1-15, 2018.
14. G. Wang, Y. Xu, S. Qin, Basic Fuzzy Event Space and Probability Distribution of Probability Fuzzy Space. Mathematics.Vol. 7, No. 6, pp. 1-15, 2019.
15. [21] P.Y. Hao, J.H. Chiang, Senior Member, IEEE. Fuzzy Regressions Analysis by Support Vector Learning Approach, Vol. 16, No. 2, pp. 428-441, April 2008.
16. P. Škrabánek, J. Marek, MODELS USED IN FUZZY LINEAR REGRESSION, 17th Conference on Applied Mathematics APLIMAT, pp. 954-964, 2018.
17. S. Sulistyo, S. Alam, SINR and throughput improvement for VANET using fuzzy power control. International Journal of Communication Systems, Vol. 31, No. 6, 2018.
18. M. Kashani, M. Ghanbari, A. Rahmani, Improving Performance of Opportunistic Routing Protocol using Fuzzy Logic for Vehicular Ad-hoc Networks in Highways. Journal of Artificial Intelligence and Data Mining. Vol. 8, No. 10, pp. 1-14, 2020.
19. C. Zhang, K. Chen, X. Zeng, and X. Xue, "Misbehavior Detection Based on Support Vector Machine and Dempster-Shafer Theory of Evidence in VANETs," in IEEE Access, vol. 6, pp. 59860-59870, 2018.
20. P. Hao, J. Chiang, "Fuzzy Regression Analysis by Support Vector Learning Approach," in IEEE Transactions on Fuzzy Systems, vol. 16, no. 2, pp. 428-441, April 2008.

21. H. Wang, J. Wang, An Effective Image Representation Method Using Kernel Classification, IEEE 26th International Conference on Tools with Artificial Intelligence pp. 853-858, 2014.
22. H. Tanaka, H. Lee, Interval Regression Analysis by Quadratic Programming Approach, IEEE.Transactions on Fuzzy Systems, Vol. 6, pp. 473-481, 1998.
23. H.C. Tsai, Unified particle swarm delivers high efficiency to particle swarm optimization, Applied Soft Computing, Vol. 55, pp. 371-383, 2017.
24. D. H. Hong, C. Hwang, Ridge regression procedures for fuzzy models using triangular fuzzy numbers, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 12, pp. 145-159, 2004.
25. B. Yelure and S. Sonavane, "Particle Swarm Optimization based Routing Method for Vehicular Ad-hoc Network," 2020 International Conference on Communication and Signal Processing (ICCSP), 2020, pp. 1573-1578.
26. M. Clerc, J. Kennedy, The particle swarm—explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. Vol. 6, pp. 58-73, (2002).
27. Q. Wu, R. Law, “Fuzzy support vector regression machine with penalizing Gaussian noises on triangular fuzzy number space”, Expert Systems with Applications, Vol. 37, No. 12, pp. 7788-7795, 2010.
28. D. Peng, M.A. Xiaoqi, W. Zhuanping, M.O. Yuanfu, P. Peng, “A prediction method of missing vehicle position information based on least square support vector machine”, Sustainable Operations and Computers, Vol. 2, pp. 30-35, 2021.