Neurocomputing 321 (2018) 61–71

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A Granular Functional Network with delay: Some dynamical properties and application to the sign prediction in social networks

Vincenzo Loia[a,*], Domenico Parente[a], Witold Pedrycz[b], Stefania Tomasiello[a]

[a] Dipartimento di Scienze Aziendali – Management & Innovation Systems (DISA-MIS), Università degli Studi di Salerno, via Giovanni Paolo II, 132, Fisciano 84084, Italy

[b] Department of Electrical Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada

ARTICLE INFO

Article history:
Received 22 May 2018
Revised 10 August 2018
Accepted 23 August 2018
Available online 4 September 2018

Communicated by Dr. Nianyin Zeng

Keywords:
Stability
Bifurcations
Neimark–Sacker
Information granularity
Fuzzy sets

ABSTRACT

In this paper, we propose a general scheme of Functional Network, by considering granularity of information and time delay. Functional Networks (FNs) are a relatively recent alternative to standard Neural Networks (NNs). They have shown better performance in comparison to performance of NNs. Data granulation used in the development of NNs allows for the formation of more efficient and transparent architectures. Time delay models have been recognized to be more realistic constructs of real-world systems. By keeping these observations in mind, we revise the usual design scheme of FN by casting it in the settings of information granules, defining a different learning algorithm, and by introducing time delay. Under some assumptions, we discuss some dynamical properties of the proposed model, in particular those concerning asymptotic stability and Neimark–Sacker bifurcation. Finally, we present an application of the proposed method to the problem of sign prediction in social networks. The results reported against those obtained by the state-of-the-art method show good performance of the proposed approach.

1. Introduction

Time delay models are intended to be more realistic constructs, since not only the current state of a system, but also the past ones are taken into account. Though there have been conflicting opinions on delay models (systems), they have been definitely accepted as a suitable way for modelling many real systems and phenomena (e.g. see [9]).

Equipped with such a vision, many works devoted to delayed neural networks, also named Time Delay Neural Networks (TDNNs), were reported. A few studies discuss some applications, e.g. [8,45], whilst there is an extensive literature devoted to the dynamical analysis (just to mention [6,20,44,46,47]), since in the design of TDNNs, the dynamical properties of the network, through stability and bifurcation analysis, play an important role. Detailed dynamical analyses of discrete cases are usually referred to particular structures, such as ring structures [27], or to a limited number of neurons [15]. The global asymptotic stability of a general discrete-time network was investigated in [19].

Here we are concerned with a discrete model, capturing the granularity of information processed by the network. By starting with the classical model of a delayed neural network (see [32] and works therein), we discretize it by processing the delayed information via granulation. Data granulation in Neural Networks was introduced to create new computing architectures, namely Granular Neural Networks (GNNs) [35], which are able to achieve a higher degree of transparency of the architecture and related operational steps. This is achieved by invoking information granules, constructs which may be expressed (formalized) through set theory, fuzzy sets, rough sets, or probabilistic information granules. Information granules support a faster learning process, thanks to their abstraction abilities. Many works devoted to granular computing have been reported in the literature (e.g., see [23,34,40]).

The model proposed here can be regarded as a discrete version of nonlinear coupled oscillators with delay. In particular, by processing the delayed data by information granules, we find a structure, which can be regarded as a generalization of the architecture proposed in [25], where the authors applied the notion of information granularity to Functional Networks (FNs). FNs were introduced by Castillo [4] as arbitrary functional models, where the activation functions are unknown multivariate functions to be determined during the learning process. As shown in [5], every problem which can be solved by means of Neural Networks (NNs) can be formulated by means of FNs and all the problems which cannot be solved by NNs can be formulated by means of FNs.

* Corresponding author.
E-mail addresses: loia@unisa.it (V. Loia), parente@unisa.it (D. Parente), wpedrycz@ualberta.ca (W. Pedrycz), stomasiello@unisa.it (S. Tomasiello).

https://doi.org/10.1016/j.neucom.2018.08.047
0925-2312/

Due to the complexity of the system, we perform a restricted stability analysis, completed under certain assumptions, in order to investigate the dynamics, as usual in such cases (e.g. [17]). We perform some numerical experiments, by starting with a two-neurons network in order to capture the dynamics, confirming the theoretical aspects of the study. Afterwards, we apply the proposed model to predict the signs of edges in social networks.

Social interaction occurring on the Web involves both positive and negative relationships: users create links to indicate friendship, support, or approval, but they also build links to express disagreement or distrust of the others' opinions. The sign prediction problem can be used to estimate the sentiment of individuals toward each other, given information about the current state of the network. Suggesting new relationships to a user, in particular with whom one shares friends, interests, or other properties, is an usual issue in online networks. Estimating the attitudes of an user from existing evidence, namely both positive and negative relationships, in the network is a challenging task.

The problem has been handled by means of logistic regression [24,37,44], probabilistic approaches [21,43], and time-delay neural networks [2]. In particular, in the pioneering work [24], the numerical experiments were performed by representing each edge of the network both as a 7-dimensional vector of degree features and as a 16-dimensional vector of triad features. More recent works devoted attention to the first or the second representation according to the kind of approach. For instance, in [2], the authors showed that the first representation ensures a high accuracy, by using a TDNN, though they considered a limited dataset. In [21], the authors considered only the second representation, being the most suitable one for their semi-probabilistic approach, achieving also a high accuracy. From a topological perspective, it is well-known the importance of the triadic closure, expressing the fact that at two different instants, a certain number of new edges have been formed through a triangle-closing operation, between two people who had previously a common neighbour. If from the one hand, the generic role of the triadic closure seems to justify the choice of a 16-dimensional vector of triad features, on the other hand, it reminds us that it motivated the formulation of simple social network measures to capture its prevalence. Among these there is the clustering coefficient, which has found significant attention in the literature, where efficient approximate formulas to compute it have been proposed (see [10] and references therein). The embeddedness of the endpoints of an edge is strictly related to their local cluster coefficients [10]. All this motivated us to consider a 7-dimensional, locally complete, degree futures vector, by taking into account the local cluster coefficients of the endpoints of the edges instead of their embeddedness (which was considered in [24]).

The local cluster coefficient of the ith node can be defined as the ratio between the number of triangles connected to the node i and the number of triples centered on i [31]. In other words, it is the fraction of pairs of i's friends that are connected to each other by edges. The clustering coefficient of a node may vary between 0 and 1, that is between the two extreme cases when none of the node's friends are friends with each other and when all of the node's friends are friends with each other. There are social theories on nodes with high clustering coefficients (see [10] and references therein).

Considering the local cluster coefficients instead of the embeddedness allows an easier dataset organization and speeding up the learning, whitout losing information on the edges.

In the numerical experiments, we tried first to check the stability of the proposed approach, by inducting some networks from publicly available datasets with a growing number of nodes, and then we computed the accuracy. A comparison against a state-of-the-art method showed the good performance of our approach.

The main contribution of this paper is to provide a different interpretation of information granules allowing for the first time a mathematical model to be formally investigated according to the theory of dynamical systems. The aim is to analyze the behaviour of the relevant solutions of the proposed model. There are numerous studies on the analysis of several types of NNs modelled as dynamical systems, but no similar one on GNNs or their variants. The application to the problem of sign prediction in social networks shows the potential of the proposed approach.

The paper is structured as follows: Section 2 introduces some theoretical foundations; in Section 3 the model is presented; Section 4 is devoted to the stability analysis; in Section 5 numerical experiments for the application example are discussed and finally Section 6 covers some conclusions.

2. Preliminaries

In this section, some basic notions are introduced. Herein we refer to data granulation completed with fuzzy sets. The principle of justifiable granularity offers a conceptual and algorithmic framework in which information granules are constructed on a basis of some numeric experimental evidence. The two intuitively appealing criteria guiding a formation of a certain information granule A are coverage and specificity. Let us recall these two concepts:

(i) coverage, we quantify how many data are "covered" by A; the higher the coverage, the more justified (legitimized by data) A is;
(ii) specificity, we require that information granule to retain its sound semantics should be as specific (detailed) as possible.

The two requirements are in conflict: increasing coverage reduces specificity and vice versa. Formally, the coverage is expressed as a cardinality of information granule; say a sum of characteristic functions of A for given data. The specificity is expressed as an inverse of the support length of the granule.

These two conflicting requirements as identified above can be expressed through the following combined performance index [35]:

$$maximize \sum_k A(z_k)/supp(A), \tag{1}$$

where A is the information granule, which may belong to a certain family of fuzzy sets, and $supp(A)$ is its support. The maximization of (1) is realized with respect to the parameters of the information granules, say the bounds of the interval information granule.

Now, we recall the notions related to fuzzy sets and useful to further reading.

Let $I = [z_1, z_m]$ be a closed interval and $z_1 < z_2 < \cdots < z_m$, with $m \geq 3$, be points (called nodes) of I. A fuzzy partition of I is defined as a family $\{A_1, A_2, \ldots, A_m\}$ of fuzzy sets A_i: $I \to [0, 1]$, $i = 1, \ldots, m$. The basic functions A_i satisfy the properties of continuity, positivity (with a unique maximum at z_i), locality (they are null outside their support) and a partition condition (the sum of membership grades sum up to one for all $z \in I$). The norm of the partition is $h = max_i |z_{i+1} - z_i|$.

A fuzzy partition can be realized by means of several basic functions. For instance, typical basic functions are the triangular ones. In order to satisfy (1), in this study we will consider fuzzy partitions with small support. A fuzzy partition with small support has the additional property that there exists $r \geq 1$ such that $supp(A_i) = \{z \in I : A_i(z) > 0\} \subseteq [z_i, z_{i+r}]$. Possible basic functions to realize such fuzzy partitions are Bernstein basis polynomials and B-splines (e.g. see [3]). An explicit form of the cubic B–splines, for

$j = 0, \ldots, m$, is given as follows (e.g. see [29])

$$A_j(z) = \frac{1}{h^3}\begin{cases} (z - z_{j-2})^3, & z \in [z_{j-2}, z_{j-1}) \\ (z - z_{j-2})^3 - 4(z - z_{j-1})^3, & z \in [z_{j-1}, z_j) \\ (z_{j+2} - z)^3 - 4(z_{j+1} - z)^3, & z \in [z_j, z_{j+1}) \\ (z_{j+2} - z)^3, & z \in [z_{j+1}, z_{j+2}) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

It should be pointed out that in order to apply B-splines, some auxiliary points are needed: for cubic B-splines two auxiliary points both on the left and on the right of the considered interval are required.

The Bernstein basis polynomials of degree p are expressed as follows

$$A_j(z) = \binom{p}{j} z^j (1 - z)^{p-j}, \quad (3)$$

$j = 0, 1, \ldots, p$, $z \in [0, 1]$.

In this case, for a certain degree p, the fuzzy partition will be given by $m = p + 1$ basic functions.

3. The proposed model

3.1. Description

It is well-known that neural networks can be regarded as nonlinear dynamical systems or nonlinear coupled oscillators modeled by the general equation [17]

$$\dot{u}_j(t) = \overline{F}_j(u_j(t)), \quad j = 1, \ldots, n, \quad (4)$$

where $u_1, \ldots, u_n$ denote the state variables of the dynamical system and $\overline{F}_j(.)$ a nonlinear function of its argument.

Here we consider the following model with delay (e.g. see [32] and papers citing it, starting with [33]), that is the system

$$\dot{u}_j(t) = -a u_j(t) + \sum_{k=1}^{n} b_{jk} \overline{f}(u_k(t - \tau_k)), \quad j = 1, \ldots, n. \quad (5)$$

The asymptotic stability of (5) when $\overline{f}(.)$ is linear and for a single delay $\tau_k = \tau$ was first investigated in [30]. Here, we assume $a \in (0, 1]$.

By fixing $0 < \tau \le 1$ as the step size of the discretization, then the discretized equations by means of forward finite differences can be written as

$$u_j^{i+1} = w_0 u_j^i + \sum_{k=1}^{n} c_{jk} \overline{f}(u_k^{i-1}), \quad j = 1, \ldots, n \quad (6)$$

where $w_0 \in [0, 1)$ takes into account the spacing size of the discretization, as well as the real coefficients c_{jk}.

Let $A_l(.)$ be m basic functions (see Section 2) giving a fuzzy partition of the reference domain, then we rewrite (6) as follows

$$u_j^{i+1} = w_0 u_j^i + \sum_{k=1}^{n} \sum_{l=1}^{m} (-1)^k A_l(u_k^{i-1}) w_{kl}^{(j)}, \quad j = 1, \ldots, n \quad (7)$$

where the weights $w_{kl}^{(j)} \in [0, 1]$ are such that $\sum_{l=1}^{m} w_{kl}^{(j)} = 1$, with $k = 1, \ldots, n$.

The resulting scheme is depicted in Fig. 1. It can be regarded as a three-layered structure:

- input layer, where the units are intended to provide the data at the time step $i - 1$ and i;
- granular layer, that is a collection of fuzzy granules, processing the input data;
- output layer, giving the required value.

Let $\{u_1^{i-1}, u_2^{i-1}, \ldots, u_n^{i-1}\}$ be the n–dimensional input at the time step $i - 1$. For any input u_k^i, the output u_k^{i+1}, with $k = 1, \ldots, n$, is obtained in additive form by processing the input at the time step $i - 1$ through the granular layer. This layer is built through some fuzzy sets A_l^k, which form a fuzzy partition of the kth input domain $\bar{I}_k$, with $k = 1, \ldots, n$ and $l = 1, \ldots, m$. Hence, for any $\bar{u} \in \bar{I}_k$, there is an associated m-tuple.

A granule is built using a certain number of fuzzy sets and it is usually intended as a relation [36], which may be formulated in several ways (e.g. see [23,26]). Here, we offer a different interpretation of the granule leading to the model (7).

Let A_l be normal and convex fuzzy sets, $l = 1, \ldots, m$. We assume that for each granule Γ^r there exists a possibility distribution $\phi_\Gamma \in \mathbb{R}^n$, such that

$$\max_{u_j, j=1,\ldots,n, j \neq r} \phi_\Gamma(u_1^{i-1}, u_2^{i-1}, \ldots, u_n^{i-1}) = A_l(u_r^{i-1}) \quad (8)$$

For each granule we carry out the following approximation:

$$f^l = f^l(A_l(u_1^{i-1}), \ldots, A_l(u_n^{i-1})) = \bigvee_{l=1}^{n} A_l(u_k^{i-1}) * w_{kl}, \quad (9)$$

where $*$ is a t-norm, $\bigvee$ is the maximum operator, $A_l(u_k^{i-1})$ is the membership degree of the data from the kth input domain and w_{kl} are the weights associated to a weighted average operator of order n, intended as a mapping $f : \mathbb{R}^n \to \mathbb{R}$, such that $w_{kl} \in [0, 1]$ and $\sum_{k=1}^{n} w_{kl} = 1$ (for the sake of simplicity, we omitted the apex (j)).

For uniformity of notation, we write $f^0 = f^0(A_0(u_j^i)) = w_0 u_j^i$, where A_0 denotes here the identity function. The output u_j^{i+1} is processed through a real invertible function F, that is $F(u_j^{i+1})$.

The small black circles shown in Fig. 1 represent the sign imposition consistent with the model (7). This sign imposition is important in addressing stability issues (see Section 4).

From the proposed scheme, one may extract rules such as:

R_j^s: IF u_1^{i-1} is A_s AND u_2^{i-1} is A_s AND ... u_n^{i-1} is A_s THEN $u_j^{i+1} = F^{-1}(f^s(A_s(u_1^{i-1}), \ldots, A_s(u_n^{i-1})) + f^0(A_0(u_j^i)))$, which are Takagi–Sugeno—Kang-like rules.

In the following, we will consider F as the identity function. A common choice for F is the linear function (e.g. see [25,41] and related references therein). Because of the adopted learning algorithm (see next section), having a multiplication term would not be significant.

3.2. The learning algorithm

The proposed network is trained through a supervised learning, by solving a minimization problem. In supervised learning, the network parameters are adjusted as a result of the influence of the training vector and the error signal, being the latter the difference between the required response and the actual response of the network. All these results in an error correction procedure.

Herein, initially we assign randomly the parameter $w_0 \in (0, 1)$. Let $\overline{N}$ be the number of sampling patterns in the dataset and $d = nm$. The unknown coefficients $w_{kl}^{(j)}$ are learnt by solving the following constrained least squares (CLS) problem

$$\min_u \|\mathbf{A}\mathbf{w} - \mathbf{y}_j\|^2 \quad (10)$$

$$s.t. \quad \mathbf{1}^T \mathbf{w}_l = 1, \quad (11)$$

$$0 \le \mathbf{w}_l \le 1, \qquad l = 1, \ldots, m \quad (12)$$

where the apex (j) was omitted for the sake of readability and with $\mathbf{y}_j^T = \mathbf{u}_j^{i+1} - w_0 \mathbf{u}_j^i$ an $\overline{N}$-dimensional vector, $\mathbf{A} \in \mathbb{R}^{\overline{N} \times d}$ the matrix of which kth row is $\{\mathbf{A}_1(\mathbf{u}_k), \mathbf{A}_2(\mathbf{u}_k), \ldots, \mathbf{A}_m(\mathbf{u}_k)\}$, being $\mathbf{u}_k$ the kth n-dimensional input pattern; $\mathbf{1}^T$ is a row vector with all

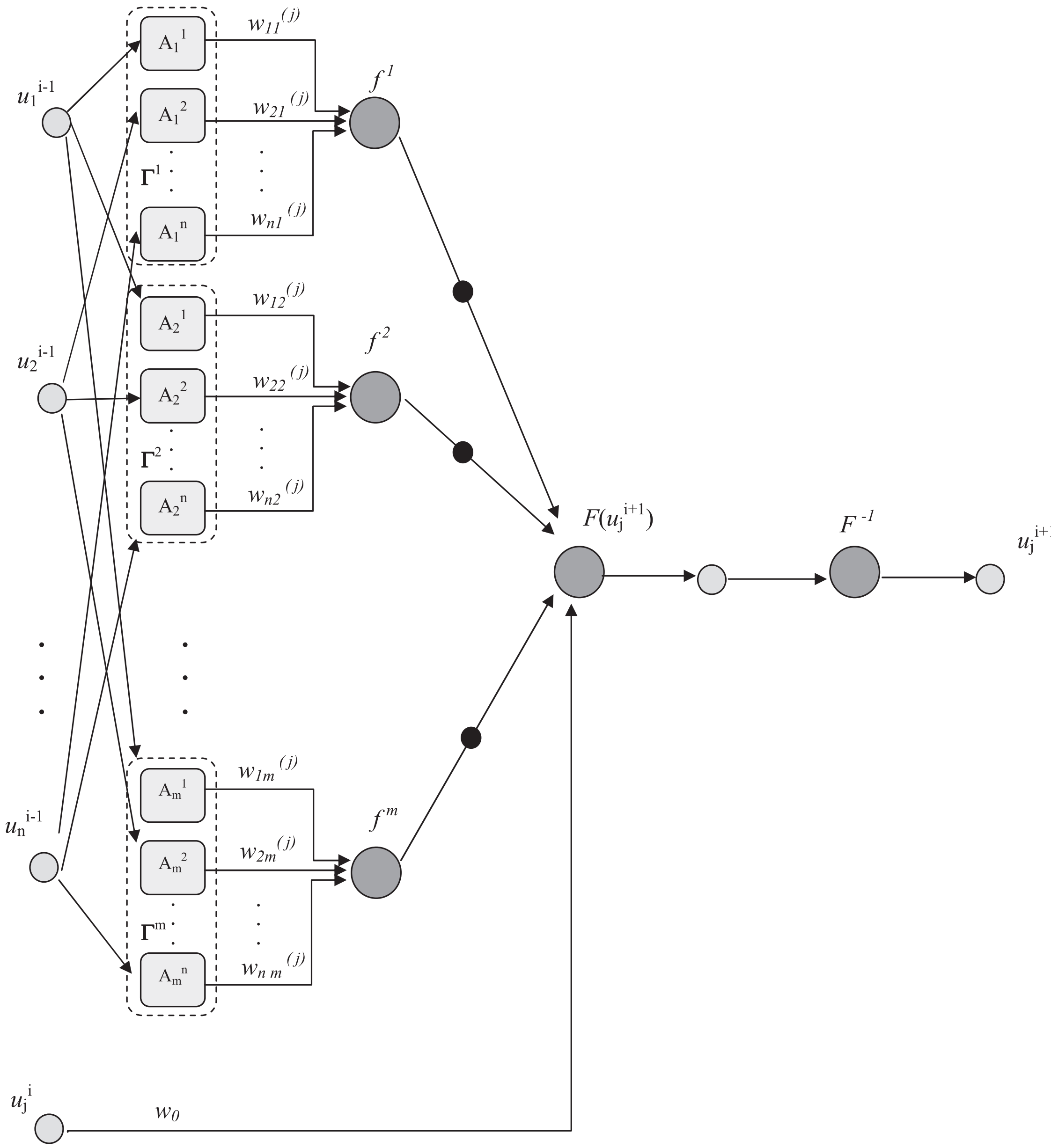

Fig. 1. The proposed model.

the elements equal to 1 and $\mathbf{w}^T = (\mathbf{w}_1^T, \ldots, \mathbf{w}_m^T)$ is the vector of unknowns (herein collected in subvectors). The lower and upper bounds of $\mathbf{w}_l$ in the CLS problem above have to be intended as processed componentwise, of course.

It should be pointed out that in the usual FNs the learning of the coefficients is also based on a CLS problem, but the constraint is given by an arbitrary initial condition, without any bound. A typical algorithm for solving the CLS problem above, where there are equality constraints as well as upper and lower bounds, is the *interior-point* method (e.g. see [42]). When the number of unknowns is significantly large, a genetic algorithm can be used to solve a CLS problem, usually by referring to a penalty function (e.g. see [38]). Hence, the computational cost of the proposed approach depends on the kind of algorithm used to solve the CLS problem. The computational cost of an interior-point method is mostly due to the solution of the linear system defining the search direction. Recently, some matrix sketching approaches have been introduced aiming at reducing the computational cost, such as the *count sketch* [7], which cost is $O(\overline{N}d\log(d))$, with $N \gg d$.

4. Restricted stability analysis

We start off this section by recalling some theoretical results present in literature.

To the end, the model (6) is rewritten in a compact form as follows

$$\mathbf{u}(i+1) = w_0\mathbf{u}(i) + \mathbf{W}\overline{\mathbf{A}}(\mathbf{u}(i-1)), \tag{13}$$

where $\mathbf{u}(i)^T = \left(u_1^i, u_2^i, \ldots, u_n^i\right)$, $\overline{\mathbf{A}}$ is the mn-dimensional vetor:

$$\overline{\mathbf{A}}^T = \left(A_1(u_1^{i-1}), A_1(u_2^{i-1}), \ldots, A_1(u_n^{i-1}), \ldots, A_m(u_1^{i-1}), \ldots, A_m(u_n^{i-1})\right),$$

$\mathbf{W}$ is the matrix whose jth row is

$$\{w_{11}^{(j)}, w_{21}^{(j)}, \ldots, w_{n1}^{(j)}, w_{12}^{(j)}, \ldots, w_{n2}^{(j)}, \ldots, w_{1m}^{(j)}, \ldots, w_{nm}^{(j)}\}.$$

(13) can be regarded as a special case of the one considered in [19], with no leakage delay, the identity function to activate $\mathbf{u}(i)$, and no external additional terms are present.

Since the functions A_i are continuous and bounded in the considered domain, then Theorem 3.1 in [19] can be adapted to state the global asymptotic stability of the proposed network model.

Anyway, in order to sketch part of the dynamics of the system, we state some theoretical results under certain assumptions.

Assumption 1. In the jth expression of (7) we assume that the contribution of any $w_{kl}^{(j)}$, with $k \notin \{j, j+1\}$ is negligible.

According to Assumption 1, we can consider the network as an ensemble of neurons pairs. Then one can investigate the conditions when a couple of neurons is destabilized with the others remaining at an equilibrium position (e.g. see [17]). On the other hand, considering the overall system as an interconnection of subsystems is an established way for a qualitative analysis of neural networks [28].

Hence, we will focus our attention on (7) with $n = 2$. In order to elucidate the geometry of the dynamics, it is convenient to convert the above mentioned equations system into an equivalent first-order system. To the end, we let

$$u_1^{i-1} = x_1^i, \qquad u_1^i = x_2^i, \tag{14}$$

$$u_2^{i-1} = x_3^i, \qquad u_2^i = x_4^i, \tag{15}$$

so that (7) can be rewritten as follows

$$x_1^{i+1} = x_2^i, \tag{16}$$

$$x_2^{i+1} = w_0 x_2^i - \sum_{l=1}^{m} A_l(x_1^i) w_{1l}^{(1)} + A_l(x_3^i) w_{2l}^{(1)}, \tag{17}$$

$$x_3^{i+1} = x_4^i \tag{18}$$

$$x_4^{i+1} = w_0 x_4^i - \sum_{l=1}^{m} A_l(x_1^i) w_{1l}^{(2)} + A_l(x_3^i) w_{2l}^{(2)}. \tag{19}$$

The system of difference equations above can be written as the iteration of the map

$$\mathbf{x} \mapsto \mathbf{g}(\mathbf{x}) \tag{20}$$

where $\mathbf{x}^T = (x_1, x_2, x_3, x_4)$ and $\mathbf{g}(\mathbf{x})^T = (x_2, w_0 x_2 - \sum_{l=1}^{m} A_l(x_1) w_{1l}^{(1)} + A_l(x_3) w_{2l}^{(1)}, x_4, w_0 x_4 - \sum_{l=1}^{m} A_l(x_1) w_{1l}^{(2)} + A_l(x_3) w_{2l}^{(2)})$.

In what follows we assume that the map (20) is C^1, with $\mathbf{g}$: $D \to D$, with $D = I_1 \times I_2 \times I_3 \times I_4$, being I_j, $j = 1, \ldots, 4$ bounded intervals in $\mathbb{R}$.

A fixed point $\bar{\mathbf{x}}$ is such that $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$. So, we state the following. For the sake of readability, but without compromising the generality of the results, we have assumed that the partitions of the intervals I_j, with $j = 1, \ldots, 4$ have all the same length m.

Lemma 2. *Let χ be a node of an arbitrary partition of I_1 and I_2 and ξ a node of an arbitrary partition of I_3 and I_4. Then the map (20) has the trivial fixed point $\bar{\mathbf{x}} = \{\chi, \chi, \xi, \xi\}$, if $-w_{1s}^{(1)} + w_{2s}^{(1)} = -w_{1s}^{(2)} + w_{2s}^{(2)} = 0$, for any $w_0 \in \mathbb{R}$ and $s \in \{1, \ldots, m\}$.*

Proof. Let $x_{j,1} < \ldots < x_{j,m}$ be a fixed (though arbitrary) partition of I_j. Since $A_s(x_{j,s}) = 1$ for any $j \in \{1, \ldots, 4\}$ and $s \in \{1, \ldots, m\}$, by considering $x_{1,s} = x_{2,s} = \chi$ and $x_{3,s} = x_{4,s} = \xi$ the conclusion readily holds. □

Lemma 3. *Suppose that there exist some values of $w_0 \in (0, 1)$ such that $\chi = \frac{-w_{1r}^{(1)} + w_{2r}^{(1)}}{1 - w_0}$ is the rth node of a given partition of I_1 and I_2 and $\xi = \frac{-w_{1s}^{(2)} + w_{2s}^{(2)}}{1 - w_0}$ is the sth node of a given partition of I_3 and I_4, with $r, s \in \{1, \ldots, m\}$. Then the map (20) has the not trivial fixed point $\bar{\mathbf{x}} = \{\chi, \chi, \xi, \xi\}$.*

Proof. The proof follows from Lemma 2. □

Let us consider the linearization of the map (20) at the fixed point $\bar{\mathbf{x}}^T = (\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4})$ (e.g. see [16])

$$\mathbf{x} \mapsto \mathbf{J}(\bar{\mathbf{x}})\mathbf{x} \tag{21}$$

where $\mathbf{J}(\mathbf{x})$ is the Jacobian matrix of which kj-entry is $\frac{\partial g_k}{\partial x_j}$.

The Jacobian matrix can be written as

$$\mathbf{J}(\bar{\mathbf{x}}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ p_1 & w_0 & p_2 & 0 \\ 0 & 0 & 0 & 1 \\ q_1 & 0 & q_2 & w_0 \end{pmatrix} \tag{22}$$

where $p_1 = -\sum_{l=1}^{m} w_{1l}^{(1)} \frac{dA_l(\bar{x}_1)}{dx_1}$, $p_2 = \sum_{l=1}^{m} w_{2l}^{(1)} \frac{dA_l(\bar{x}_3)}{dx_3}$, $q_1 = -\sum_{l=1}^{m} w_{1l}^{(2)} \frac{dA_l(\bar{x}_1)}{dx_1}$, $q_2 = \sum_{l=1}^{m} w_{2l}^{(2)} \frac{dA_l(\bar{x}_3)}{dx_3}$.

The related characteristic equation reads as

$$\lambda^4 - 2w_0\lambda^3 + (w_0^2 - p_1 - q_2)\lambda^2 + w_0(p_1 + q_2)\lambda - b = 0 \tag{23}$$

with $b = p_2 q_1 - p_1 q_2$.

Let us introduce the next notation

$$b_0 = -\left[w_0^4 + 4w_0^2 + 1 - (4w_0^3 + 4w_0)\cos(\phi) + 2w_0^2\cos(2\phi)\right]^{1/2},$$

where $0 < \phi < \frac{\Pi}{4}$ is the solution of the equation $\sin(4\theta) - 2w_0\sin(3\theta) + w_0^2\sin(2\theta) = 0$.

Theorem 4. *(a) Let $w_0 \in (0, 1)$. If $p_1 = -q_2$, then the trivial fixed point $\bar{\mathbf{x}}$*

- *is asymptotically stable for $b_0 < b < (1 - w_0)^2$,*
- *presents a Neimark–Sacker bifurcation for $b = -b_0$;*
- *presents a fold bifurcation for $b = (1 - w_0)^2$.*

(b) Let $w_0 = 0$. If $p_2 q_1 \geq 0$ and $\max\{|p_1| + |p_2|, |q_1| + |q_2|\} < 1$, then the trivial fixed point $\bar{\mathbf{x}}$ is asymptotically stable.

Proof. We prove first part (a). If $p_1 = -q_2$, then (23) reduces to

$$\lambda^4 - 2w_0\lambda^3 + w_0^2\lambda^2 - b = 0. \tag{24}$$

For $w_0 \in (0, 1)$, (24) was investigated in [12] and part (a) recalls the main result in [12].

With regard to part (b), when $w_0 = 0$, then (23) reduces to

$$\zeta^2 - \tau\zeta + \delta = 0, \tag{25}$$

where $\zeta = \lambda^2$ and $\tau = p_1 + q_2$, $\delta = p_1 q_2 - q_1 p_2$ are the trace and the determinant respectively of the matrix $\mathbf{M}$

$$\mathbf{M} = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}. \tag{26}$$

It is easy to check that the eigenvalues of the matrix (26) are always real when $p_2 q_1 > 0$.

Besides, we recall that for any norm $\|.\|$, it is $\rho(\mathbf{M}) \leq \|M\|$, being $\rho(\mathbf{M})$ the spectral radius of $\mathbf{M}$.

Hence, since $\|\mathbf{M}\|_\infty = \max\{|p_1| + |p_2|, |q_1| + |q_2|\}$, the conclusion follows. □

It is the case to recall that the Neimark–Sacker bifurcation is also known as Poincarè–Andronov–Hopf bifurcation for maps [16] or Hopf bifurcation for maps, for short.

4.1. An example: a two-neuron network

Before discussing the application case, it is useful to show the dynamics by means of a simple example. This is an usual way to provide an easy visualization of the dynamical behaviour of the

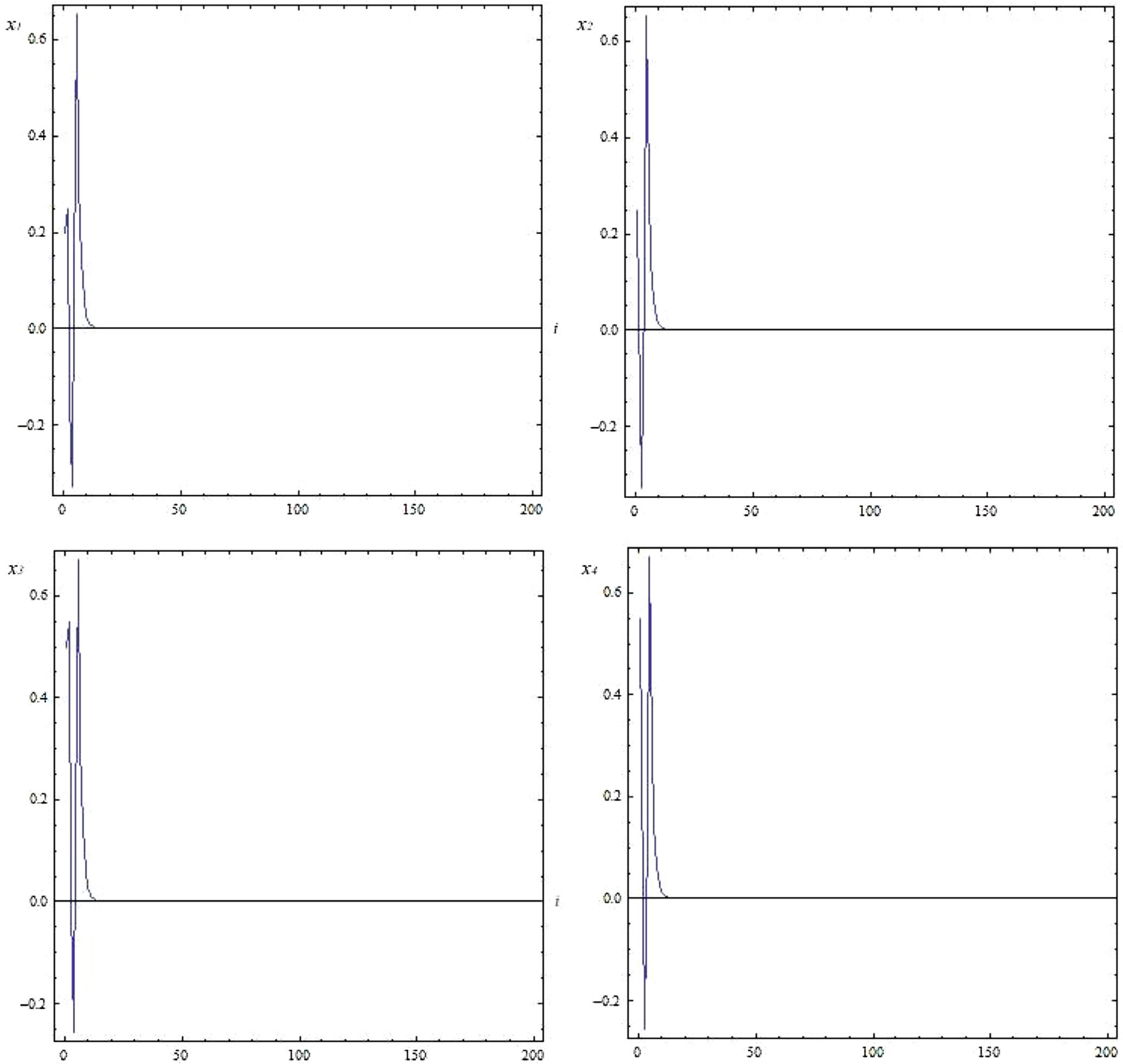

Fig. 2. A two-neuron network, first case: x_1, x_2, x_3, x_4 trajectories.

system over the theoretical results (e.g. [15]). To the end, we consider the following network, by using Bernstein basis polynomials with $m = 6$:

$$u_j^{i+1} = w_0 u_j^i + \sum_{r=1}^{2} (-1)^r [w_{r1}^{(j)} \left(1 - u_r^{i-1}\right)^5 + 5 w_{r2}^{(j)} \left(1 - u_r^{i-1}\right)^4 u_r^{i-1} + 10 w_{r3}^{(j)} \left(1 - u_r^{i-1}\right)^3 \left(u_r^{i-1}\right)^2], \tag{27}$$

with $r = 1, 2$.

As a first case, we fix $w_0 = 0.49$, $w_{11}^{(j)} = w_{21}^{(j)} = 0.45$, $w_{12}^{(j)} = w_{22}^{(j)} = 0.55$, for $j = 1, 2$, while the other coefficients are settled to zero. In this way, $b = 0$ and we are under the hypotheses of Lemma 1 and the first case of Theorem 1(a), that is there exists a trivial fixed point which is asymptotically stable. By arranging (27) into the form (19) and by assuming the initial conditions $(x_1^0, x_2^0, x_3^0, x_4^0)^T = (0.2, 0.25, 0.5, 0.55)$, it can be easily seen from Fig. 2 that the trajectories go to zero.

Now, let $w_{11}^{(j)} = w_{21}^{(j)} = 0.45$, $w_{12}^{(j)} = 0.35$, $w_{22}^{(j)} = 0.57$. We choose w_0 as bifurcation parameter: we vary it, by lefting unchanged the other coefficients. Fig. 3 shows the effect of the variation of w_0 on the stability: the typical Neimark–Sacker bifurcation phenomenon appears, by starting with $w_0 = 0.9997$, and which ends up with an attracting closed curve encircling the trivial fixed point by increasing w_0 (x_4 vs x_3 plots are very similar and not reported here for the sake of brevity). All the plots were obtained by using 1200 points.

Finally, we fix $w_0 = 0$, $w_{12}^{(1)} = w_{22}^{(2)} = 0.1$, $w_{13}^{(1)} = w_{23}^{(2)} = 0.9$, with all the other coefficients set to zero. It is easy to see that we are under the hypotheses of Theorem 1(b). Due to the coefficients choice, the system reduces to two uncoupled oscillators. The

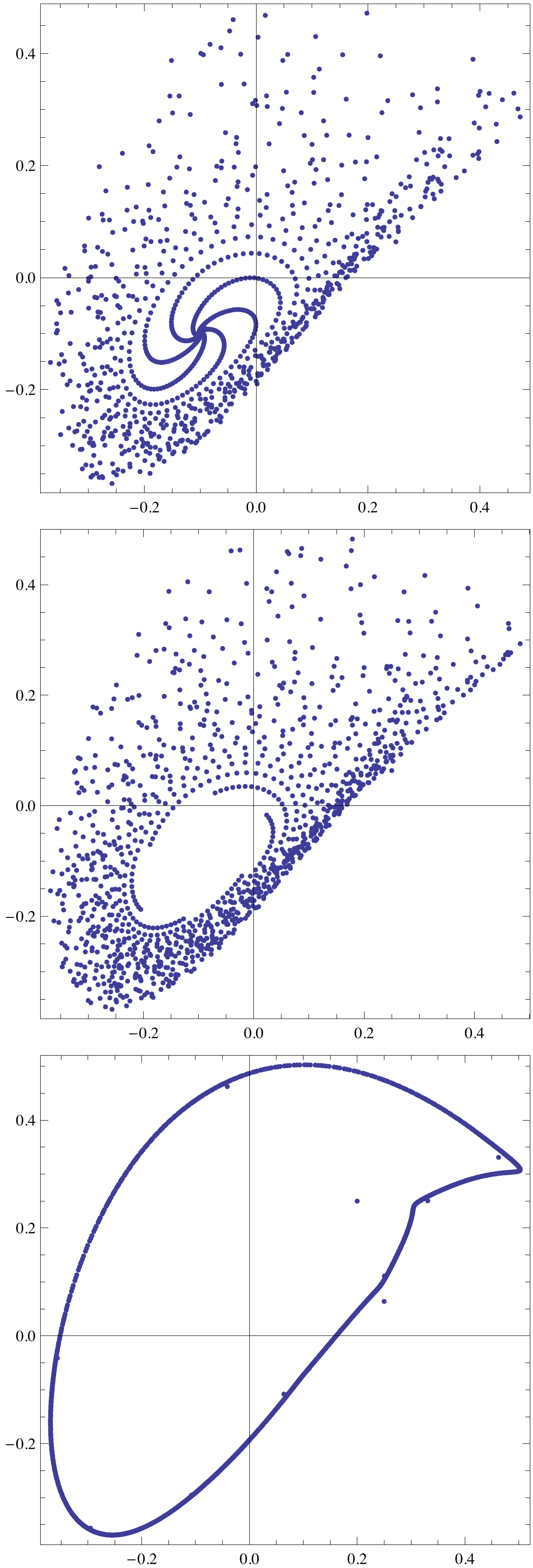

Fig. 3. A two-neuron network, second case: x_2 vs x_1 plot. From the top, $w_0 = 0.9997$, $w_0 = 0.9998$, $w_0 = 1$.

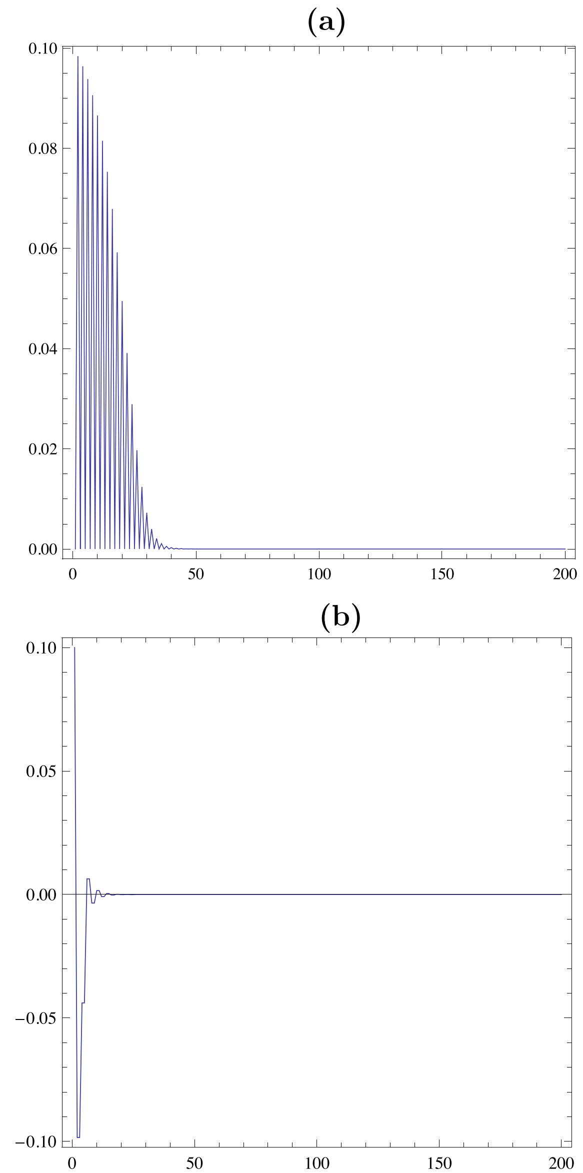

Fig. 4. A two-neuron network, third case: (a) first neuron trajectory (b) second neuron trajectory.

trajectories of these two oscillators, with the same initial condition $x_j^0 = 0.1$, are shown in Fig. 4.

5. Application to social networks

In this section, we describe the application example. We start off, by formulating the problem, then we give a brief description of the datasets used for the numerical experiments. Social networks are nowadays an attractive research topic and many related issues have been tackled by means of several types of NNs, such as cascade outbreak prediction [13], malicious user detection [18] or sign edges prediction [2]. The latter is the problem herein considered. We recall that many approaches [2,24] have been proposed for such problem, by considering a different number of features for modelling the problem. In [2], the authors showed that using 7 features in a TDNN ensures a high accuracy against other methods. In what follows, we consider the same number of features, though partially different as detailed below, by chosing as benchmark method the TDNN.

5.1. Problem formulation

Let $G = (V, E)$ be a directed graph with a sign (positive or negative) on each edge and let $s(e, f)$ denote the sign of the edge (e, f) from e to f. It is assumed $s(e, f) \in \{-1, 0, 1\}$ to indicate, in the order, a negative, null or positive sign of (e, f); the null sign applies when there is no directed edge from e to f. The notation $\bar{s}(e, f) = 1$ is used when there is a positive edge in one of the directions (e, f) or (f, e), and either a positive edge or no edge in the other direction. Similarly, it is $\bar{s}(e, f) = -1$ when there is a negative edge in one of the mentioned directions, and either a negative edge or no edge in the other direction. It is assumed $\bar{s}(e, f) = 0$ in all other cases.

5.2. Dataset description

As in [2,24], we consider the Slashdot and the Epinion datasets [39]. Slashdot is a website reporting science and technology news. On this website, users can tag each other as friend or foe generating positive or negative edges, respectively. A positive sign means that a user likes the opinions of another user, while a negative sign implies that the user dislikes or does not find trustworthy the other user's opinions. The dataset contains 82,140 nodes and 549,202 edges of which 77.4% are positive. Epinions is a customer review website, where users are connected as trusty or untrustworthy according to the review ratings. The network has 119,217 nodes and 841,000 edges, of which 85.0% are positive.

As in [2] and formerly in [24], we perform our experiments by using 7 features, based on the degrees of the nodes, which substantially summarizes the aggregate local relations of a node with respect to the rest of the network. Since the aim is predicting the sign of the edge from e to f, we consider outgoing edges from e and incoming edges to f. Then, we will use the following entities, of which just one (the ratio C) is different from the ones considered in [24]:

- $d_{in}^{+}(f)$, the number of incoming positive edges to f;
- $d_{in}^{-}(f)$, the number of incoming negative edges to f;
- $d_{out}^{+}(e)$, the number of outgoing positive edges from e;
- $d_{out}^{-}(e)$, the number of outgoing negative edges from e;
- $C = C_e/C_f$, the ratio between the local cluster coefficient of e and f, respectively, being the local cluster coefficient the ratio of the total number of existing connections between all the node's nearest neighbours k and the number $k(k-1)/2$;
- $d_{out}^{+}(e) + d_{out}^{-}(e)$, the total out-degree of e;
- $d_{in}^{+}(f) + d_{in}^{-}(f)$, the total in-degree of f.

5.3. Numerical experiments

We started numerical experiments as in [2], by considering an induced subnetwork with 2000 nodes with almost the same percentage of positive and negative signs as in the whole network. Then, in order to show the stability of the approach, we considered induced subnetworks with 4,000 and 8,000 nodes. On the other hand, considering induced subnetworks is consistent with the current research on social networks analysis, which devotes a certain attention to sub-graphs of the network graph (e.g. [1–22]).

Here, we are concerned with a one-step ahead prediction. We used a 5-fold cross validation. We performed experiments by considering both Bernstein basis polynomials and cubic B-splines as basic functions. Data were normalized in the range [0,1]. For comparison purposes, we chose as benchmark method the TDNN, since in [2] it was shown to be the best choice against some other methods, especially when using 7 feautures for modelling the problem. In such a case, the results obtained in [2] were competitive (though on a scaled network) if compared with the ones referred

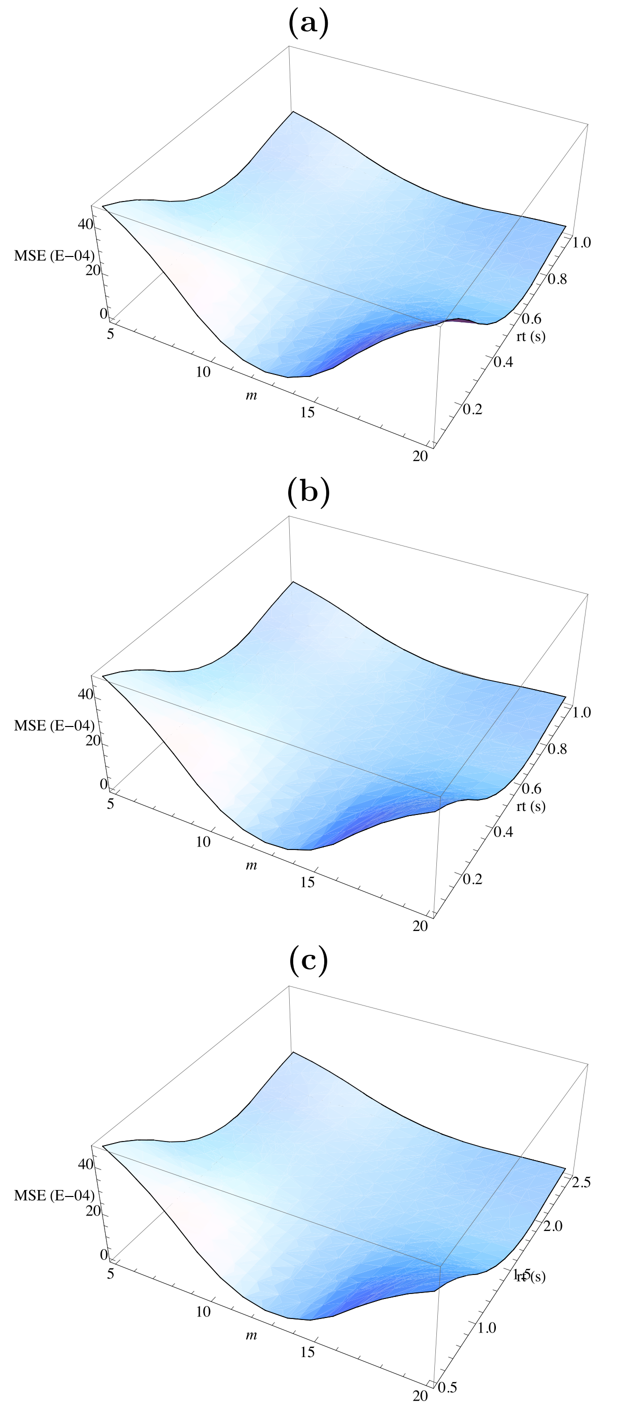

Fig. 5. MSE vs {m,rt}, cubic B-spline (Slashdot): (a) 2000 nodes (b) 4000 nodes (c) 8000 nodes.

to 16 features presented in [21]. Hence, our results were compared with the ones by a TDNN with a tapped delay 1: 8 and a hidden layer with 15 units, as adopted in [2].

As quality indicator, we considered the accuracy

$$ACC = \frac{TP + TN}{P + N}, \tag{28}$$

where

- TP (true positive) is the number of positive edges predicted correctly.
- TN (true negative), is the number of negative edges predicted correctly.

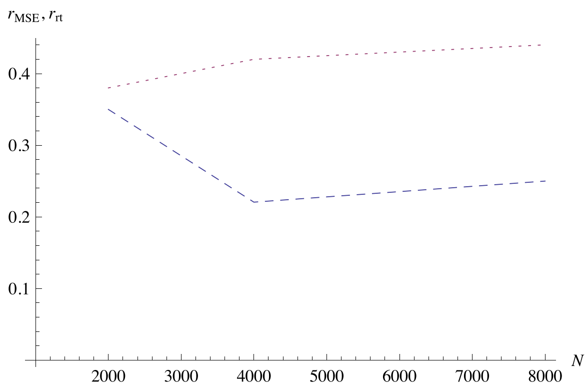

Fig. 6. *Slashdot*: r_{MSE}, r_{rt} vs N (r_{MSE}, dashed line; r_{rt}, dotted line).

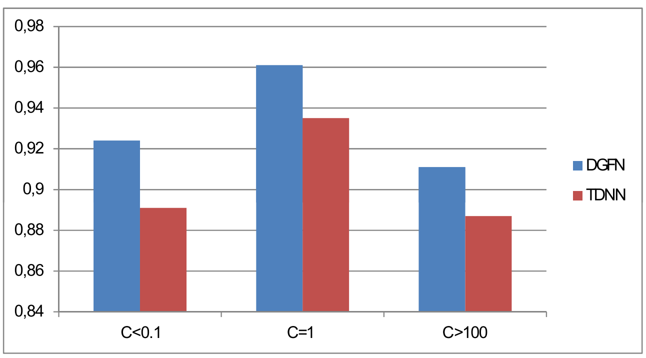

Fig. 7. Slashdot: ACC values.

- P and N, are the number of positive and negative edges respectively in the network.

In particular, we check the accuracy with regard to the three cases, namely:

- $C < 0.1$, meaning a small C_e with respect to C_f,
- $C = 1$, that is $C_e = C_f$,
- $C > 100$, meaning a small C_f with respect to C_e.

With regard to the training stage, we considered the usual Mean Squared Error (MSE). The results were compared with the one by a TDNN with a tapped delay 1:8, as adopted in [2]. To the end, we considered the ratios $r_{MSE} = MSE_{DGFN}/MSE_{TDNN}$ and $r_{rt} = rt_{DGFN}/rt_{TDNN}$, referred to the *MSE* and the running time (*rt*) in the proposed Delayed Granular Functional Network (DGFN) and the TDNN.

The numerical computations were performed by using a CPU clocking in at 2.40 GHz. The accuracy of the obtained results by our approach is consistent with that of the state-of-the-art techniques [2,21], presenting the highest accuracy against the other existing approaches, even though we do not refer to embeddedness. We also performed a few numerical experiments by considering embeddedness, but we found worser results (MSE in mean 10 times higher and higher running times).

Slashdot

In Fig. 5, the resulting MSE for the training stage is shown for the induced subnetworks with 2000, 4000 and 8000 nodes. The graphs show the MSE vs the basis function order *m* and the running time (*rt*). These results were obtained by using cubic B-splines. By using Bernstein basis polynomials, the errors are almost doubled, though the *rt* is almost the same. The *MSE* does not grow, as the network grows. The dimension of the network partially affects the *rt*, with a clearer evidence for the Epinions dataset case. The best performance is achieved for $m = 12$, with an *MSE* with order 10^{-5}, in all the induced subnetworks. Fig. 6 shows the behaviour of the ratios r_{MSE}, r_{rt} with the number of nodes N of the network, for a comparison with the TDNN. The graph was obtained by considering the best performance. As one can notice, the running times and the MSE are significantly lesser.

In Fig. 7 the average accuracy of prediction for DGFN and TDNN is shown.

Epinions

In Fig. 8 the MSE vs $\{m, rt\}$ for the training stage is shown for the induced subnetworks with 2000, 4000 and 8000 nodes. These results were obtained by using Bernstein basis polynomials. In this case, by using cubic B-splines the errors are 10 times higher. The

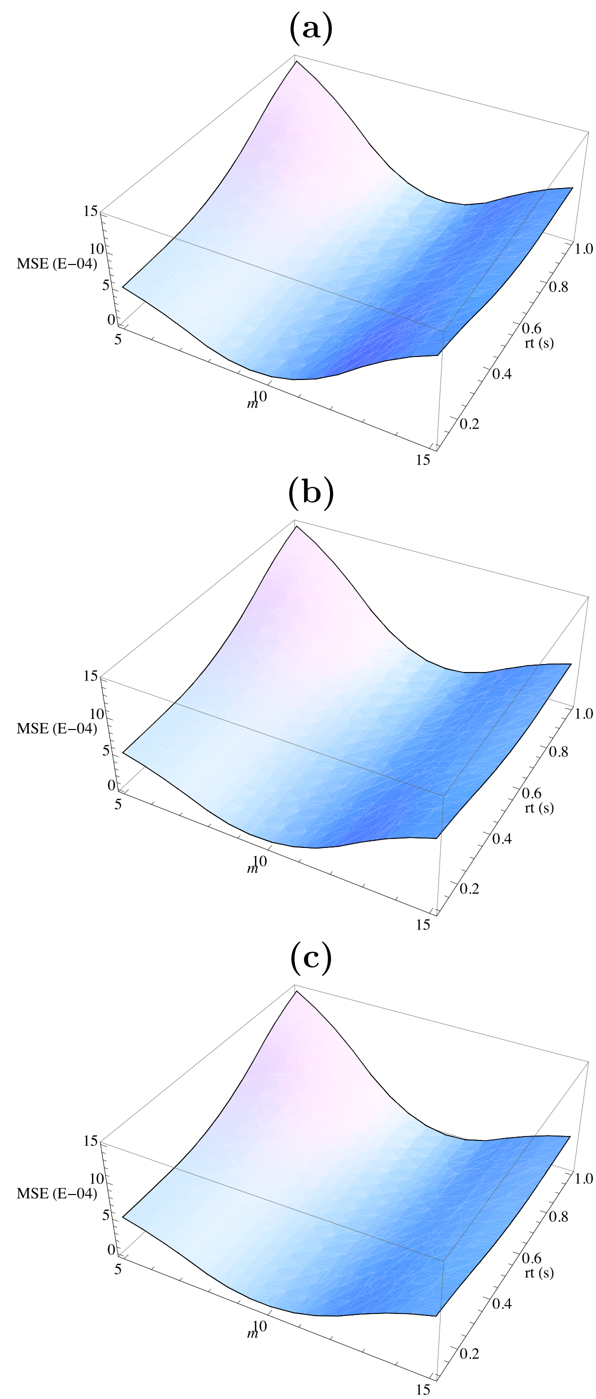

Fig. 8. MSE vs {m,rt}, Bernstein (Epinions): (a) 2000 nodes (b) 4000 nodes (c) 8000 nodes.

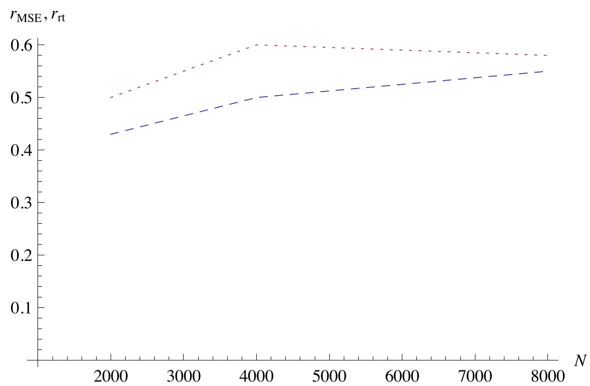

Fig. 9. Epinions: r_{MSE}, r_{rt} vs N (r_{MSE}, dashed line; r_{rt}, dotted line).

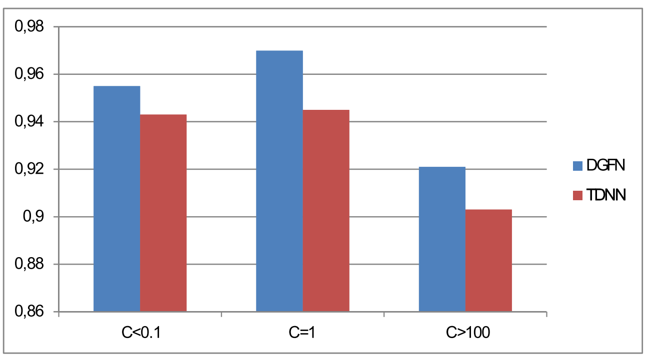

Fig. 10. Epinions: ACC values.

best performance is achieved by $m = 9$, with an MSE with order 10^{-5}, in any case.

In Fig. 9, the ratios r_{MSE}, r_{rt} vs the number of nodes N are shown. Again, the graph was obtained by considering the best performance and the MSE and running times of our approach turn out to be lesser.

In Fig. 10 the average accuracy of prediction for DGFN and TDNN is shown. By means of the DGFN the accuracy turns out to be better, with some little differences through the three cases; for instance, one can notice that for the case $C = 1$ we produced the best results.

6. Conclusions

In this paper, we investigated a different scheme of Functional Network through information granulation and time delay. We used fuzzy partitions with small support of the input data domain, introducing a granular layer into the usual scheme of a Functional Network. We defined a new learning algorithm as a constrained least square problem. Some dynamical properties of the proposed model were discussed under certain assumptions. Some numerical examples confirmed the theoretical achievements, giving a visualization of the asymptotical stability and the Newmark–Sacker bifurcation. Finally, we applied the proposed method to the problem of sign prediction in social networks. We performed some numerical experiments by using publicly available data. A comparison against a state-of-the-art method showed the efficiency of our approach. The proposed approach has the advantages of the granular network architectures, which provide a more transparent and computationally efficient structures, additionally it offers a general mathematical model, still missing in the field, bridging the gap between the classical interpretation of a network and its granular (in the fuzzy meaning) counterpart. This allows a better understanding of the behaviour of such computing architectures. In particular, the presence of a delay in the proposed model makes it a valid alternative to classical Time Delay Neural Networks.

References

[1] N.M. Ahmed, L. Chen, An efficient algorithm for link prediction in temporal uncertain social networks, Inf. Sci. 331 (2016) 120–136.
[2] S. Babakhanbak, K. Kavousi, F. Farokhi, Application of a time delay neural network for predicting positive and negative links in social networks, Turk. J. Elec. Eng. Comput. Sci. 24 (2016) 2825–2837.
[3] B. Bede, I.J. Rudas, Approximation properties of fuzzy transforms, Fuzzy Sets Syst. 180 (2011) 20–40.
[4] E. Castillo, Functional networks, Neural Proces. Lett. 7 (1998) 151–159.
[5] E. Castillo, A. Iglesias, R. Ruiz-Cobo, Functional Equations in Applied Sciences, Elsevier, The Netherlands, 2005.
[6] T. Chen, Global convergence of delayed dynamical systems, IEEE Trans. Neural Netw. 12 (6) (2001) 1532–1536.
[7] K.L. Clarkson, D.P. Woodruff, Low rank approximation and regression in input sparsity time, in: Proceedings of the Forty-fifth Annual ACM Symposium Theory Computing, ACM, 2013, pp. 81–90.
[8] J.M.M.d. Silva, E. Kaszkurewicz, A proposed solution for the load balancing problem on heterogeneous clusters based on a delayed neural network, Int. J. Intell. Comput. Cybern. 3 (1) (2010) 73–93.
[9] T. Erneux, Applied Delay Differential Equations, Springer Science, New York, NY, 2009.
[10] D. Easley, J. Kleinberg, Networks, Crowds, and Markets: Reasoning about a Highly Connected World, Cambridge University Press, UK, 2010.
[11] M. Fang, J. Yin, X. Zhu, Active exploration: simultaneous sampling and labeling for large graphs, Proceedings of the Twenty-second ACM International Conference on Information and Knowledge Management(CIKM'13) (2013) 829–834.
[12] Q. Gan, R. Xu, W. Hu, P. Yang, Bifurcation analysis for a tri-neuron discrete-time BAM neural network with delays, chaos, Solitons Fract. 42 (2009) 2502–2511.
[13] C. Gou, et al., Learning sequential features for cascade outbreak prediction, Knowl. Inf. Syst. 57 (3) (2018) 721–739.
[14] J. Guckenhelmer, P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields, Springer-Verlag, Berlin, 1983.
[15] S. Guo, X. Tang, L. Huang, Bifurcation analysis in a discrete-time single-directional network with delays, Neurocomputing 71 (2008) 1422–1435.
[16] J. Hale, H. Kocak, Dynamics and Bifurcations, Springer-Verlag, New York, 1991.
[17] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall, New Jersey, 1999.
[18] T. Hong, C. Choi, J. Shin, CNN-based malicious user detection in social networks, Concurr. Comput. Pract. Exper. 30 (2018) 41–63.
[19] L.J. Banu, P. Balasubramaniam, Robust stability analysis for discrete-time neural networks with time-varying leakage delays and random parameter uncertainties, Neurocomputing 179 (2016) 126–134.
[20] M. Joy, Results concerning the absolute stability of delayed neural networks, Neural Netw. 13 (2000) 613–616.
[21] A. Khodadadi, M. Jalili, Sign prediction in social networks based on tendency rate of equivalent micro-structures, Neurocomputing 257 (2017) 175–184.
[22] S. Kim, K. Kim, S. Yoon, S. Park, Sampling in online social networks, in: Proceedings of the Twenty-ninth Annual ACM Symposium on Applied Computing Computing (SAC'14), 2014, pp. 845–849.
[23] D. Leite, P. Costa, F. Gomide, Evolving granular neural networks from fuzzy data streams, Neural Netw. 38 (2013) 1–16.
[24] J. Leskovec, D. Huttenlocher, J. Kleinberg, Predicting positive and negative links in online social networks, in: Proceedings of the Nineteenth International Conference World Wide Web (WWW '10), 2010, pp. 641–650.
[25] V. Loia, S. Tomasiello, Granularity into functional networks, in: Proceedings of the Third IEEE International Conference on Cybernetics (CYBCONF 2017), IEEE, 2017, pp. 1–6.
[26] W. Lu, W. Pedrycz, X. Liu, J. Yang, P. Li, The modeling of time series based on fuzzy information granules, Expert Syst. Appl. 41 (2014) 3799–3808.
[27] R. Mazrooei-Sebdani, S. Farjami, On a discrete-time-delayed Hopfield neural network with ring structures and different internal decays: Bifurcations analysis and chaotic behavior, Neurocomputing 151 (P1) (2015) 188–195.
[28] A.N. Michel, J.A. Farrell, W. Porod, Qualitative analysis of neural networks, IEEE Trans. Circ. Syst. 36 (2) (1989) 229–243.
[29] R.C. Mittal, A. Tripathi, Numerical solutions of generalized Burgers–Fisher and generalized Burgers–Huxley equations using collocation of cubic B-splines, Int. J. Comput. Math. 92 (5) (2015) 745–758.
[30] T. Mori, Criteria for asymptotic stability of linear time-delay systems, IEEE Trans. Autom. Control AC-30 (2) (1985) 158–161.
[31] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.
[32] L. Olien, J. Berlair, Bifurcations, stability, and monotonicity properties of a delayed neural network model, Physica D 102 (1997) 349–363.
[33] K. Pakdaman, C. Grotta-Ragazzo, C.P. Malta, O. Arino, J.F. Vibert, Effect of delay on the boundary of the basin of attraction in a system of two neurons, Neural Netw. 11 (3) (1998) 509–519.
[34] W. Pedrycz, H.S. Park, S.K. Oh, A granular-oriented development of functional radial basis function neural networks, Neurocomputing 72 (2008) 420–435.
[35] W. Pedrycz, W. Vukovich, Granular neural networks, Neurocomputing 36 (2001a) 205–224.

[36] W. Pedrycz, G. Vukovich, Abstraction and specialization of information granules, IEEE Trans. Syst. Man Cybern. B 31 (1) (2001b) 106–111.
[37] M. Shahriari, M. Jalili, Ranking nodes in signed social networks, Soc. Netw. Anal. Min. 4 (1) (2014) 1–12.
[38] A.E. Smith, D.W. Coit, Constraint handling techniques-penalty functions, Handbook of Evolutionary Computation, Oxford University Press and Institute of Physics Publishing, UK, 1997. Ch. C 5.2
[39] http://snap.stanford.edu/data.
[40] M. Song, W. Pedrycz, Granular neural networks: concepts and development schemes, IEEE Trans. Neural Netw. 24 (2013) 542–553.
[41] S. Tomasiello, A functional network to predict fresh and hardened properties of selfcompacting concretes, Int. J. Num. Meth. Biomed. Eng. 27 (6) (2011) 840–847.
[42] R.J. Vanderbei, T.J. Carpenter, Symmetric indefinite systems for interior point methods, Math. Program. 58 (1993) 1–32.
[43] Y. Wang, F. Liu, S.T. Xia, J. Wu, Link sign prediction by variational bayesian probabilistic matrix factorization with student-t prior, Inf. Sci. 405 (2017) 175–189.
[44] W. Yuan, C. Li, G. Han, D. Guan, L. Zhou, K. He, Negative sign prediction for signed social networks, Future Gen. Comput. Syst. (2018). in press
[45] Y. Zhang, et al., A time delay neural network model for simulating eye gaze data, J. Experim. Theor. Artif. Intell. 23 (1) (2011) 111–126.
[46] Y. Zhao, H. Gao, S. Mou, Asymptotic stability analysis of neural networks with successive time delay components, Neurocomputing 71 (2008) 2848–2856.
[47] J. Zhou, Z. Liu, G. Chen, Dynamics of periodic delayed neural networks, Neural Netw. 17 (2004) 87–101.

Vincenzo Loia (IEEE SM'06) received the master's degree in computer science from the University of Salerno, Fisciano, Italy, in 1985, and the Ph.D. degree in computer science from the University of Paris 6, Paris, France, in 1989. He is a Full Professor of Computer Science in University of Salerno, where he is the Director of the Dept. of Management and Innovation Systems. He is the Editor-in-Chief of Ambient Intelligence and Humanized Computing and of Evolutionary Intelligence (both from Springer). He serves as an Associate Editor in several IEEE Transactions, among them Industrial informatics, Fuzzy Systems, Cognitive and Developmental Systems. His research interest are essentially in the area of computational intelligence, data mining, fuzzy and granular computing, knowledge discovery, adaptive and learning systems, situational awareness and cognitive cyber security. He has published numerous papers in the above areas.

Domenico Parente is a full professor of Computer Science since 2005. He is the leader of the GandALF laboratory and the co-founder of the homonym international conference. His main research interests are in the use of formal tools for the Automatic Verification of CyberPhysical Systems and the Information extraction for profiling users in big/medium data scenarios like social networks and smart environments.

Witold Pedrycz (IEEE F'98) received the M.Sc., Ph.D., and D.Sc. degrees from the Silesian University of Technology, Gliwice, Poland, in 1977, 1980, and 1984, respectively. He is a Professor and the Canada Research Chair of computational intelligence with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland. He has authored 15 research monographs covering various aspects of computational intelligence, data mining, and software engineering. His current research interests include computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data mining, fuzzy control, pattern recognition, knowledge-based neural networks, relational computing, and software engineering. He has published numerous papers in the above areas. He was a recipient of the prestigious Norbert Wiener Award from the IEEE Systems, Man, and Cybernetics Society in 2007, the IEEE Canada Computer Engineering Medal, the Cajastur Prize for Soft Computing from the European Centre for Soft Computing, the Killam Prize, and the Fuzzy Pioneer Award from the IEEE Computational Intelligence Society. He is intensively involved in editorial activities. He is the Editor-in-Chief of Information Sciences, WIREs Data Mining and Knowledge Discovery (Wiley), and the International Journal of Granular Computing (Springer). He currently serves on the Advisory Board of the IEEE Transactions on Fuzzy Systems. He is a member of a number of editorial boards of other international journals. In 2009, he was elected as a Foreign Member of the Polish Academy of Sciences. In 2012, he was elected as a fellow of the Royal Society of Canada. He has been a member of numerous program committees of the IEEE conferences in the area of fuzzy sets and neurocomputing.

Stefania Tomasiello, Ph.D. in computer science (University of Salerno, Italy), is a permanent researcher at CORISA (Research Consortium on Agent Systems), University of Salerno, in temporary leaving as Senior Research Fellow at the Dept. of Management and Innovation Systems, University of Salerno. Workpackage leader in several funded projects, she is an expert evaluator (ex-ante and ex-post) of research projects joining academia and industry for the Italian Ministry of Economic Development. She got the Italian habilitation as associate professor in Computer Science and Numerical Analysis. She has been an adjunct professor of Fundamentals of Computer Science, Human-Computer Interaction, Computational Methods and Finite Element Analysis at the University of Basilicata, Italy. TPC member in many international conferences, here included IEEE and ACM sponsored events. Her research interests lie in scientific and soft computing, fuzzy mathematics, nonlinear dynamics. She authored and co-authored numerous papers in the above mentioned areas. She is associate editor of International Journal of Computer Mathematics (Taylor&Francis), International Journal of System Assurance Engineering and Management (Springer) and editorial board member of some international journals.