Neurocomputing 168 (2015) 1181–1187

Contents lists available at ScienceDirect

Neurocomputing

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

Modeling sustainability report scoring sequences using an attractor network

Mario González [a,b], María del Mar Alonso-Almeida [c], Cassio Avila [d], David Dominguez [d,e,*]

[a] FCA, Escuela Politécnica Nacional, Quito, Ecuador
[b] FACI, Universidad Estatal de Milagro, Milagro, Guayas, Ecuador
[c] FCEE, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[d] Instituto de Fisica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[e] EPS, Universidad Autónoma de Madrid, Spain

ARTICLE INFO

Article history:
Received 8 August 2014
Received in revised form
25 April 2015
Accepted 5 May 2015
Communicated by J. Torres
Available online 14 May 2015

Keywords:
Global reporting initiative
Attractor network modeling
Metric connectivity
Mutual Information

ABSTRACT

This work experimentally explores the metric Attractor Neural Network for modeling Corporate Sustainability Reporting patterns of a set of global companies. A small-world topology configuration is used for the metric network, and compared with a configuration obtained from the Mutual Information (MI) between companies, in terms of the usual dilution and shortcut ratios. The resulting MI topology configuration is depicted for mesoscopic blocks distributed by continents and economic sectors. The reporting sequence is learned as static patterns, as well as, a temporal sequence from year 1999 to 2013. The retrieval of the sequence showed a saturation point around 2010 where the reporting pattern stalled. We showed that the MI topology configuration obtained for continents, reinforces previous research about the role of Europe as a driver about sustainability and its influence worldwide. Also, the MI configuration outlines recent (post-crises) behavior, of the involved economic sectors.

1. Introduction

Nowadays, Corporate Sustainability Reporting (CSR) emphasizes the need for companies developing a balance called triple bottom way: be sustainable in economic, environmental and social issues. In other words, companies should explain and disclose what they are doing to stay in the market, to change to a cleaner production and improve environmental policies, and to engage in activities with the development of local communities. The globalization of markets and latter financial crisis have stressed that "it is required a harmonized, standardized and objective reports from firms worldwide to understand what companies are doing and to facilitate comparisons across companies" [1]. Therefore, a standardized sustainability report could be a tool to communicate with internal and external stakeholders worldwide, a framework to assess efforts of the company in CSR and a source of public information [2,3].

The G3 version of the Global Reporting Initiatives (GRI) framework has become the *de-facto* standard for corporate sustainability under which an increasing number of businesses have been producing their reports [4]. The GRI guidelines have the potential to significantly improve the usefulness and quality of information reported by companies about their environmental, social and economic impacts and performance [5–7]. Sustainability reporting also encourages the companies to keep track of their social responsibility and sustainable development performance. Under the third generation of the GRI framework for CSR documents, also referred to as G3 framework, a CSR report is broken down into the following performance sections: Economical, Environmental, Human Rights, Labor, Product Responsibility and Society. Reporting entities shall follow the guidelines of each section and indicate the details of their performance during the past reporting period [8–11].

Here, we propose a metric Attractor Neural Network (ANN) model to predict the sustainability reporting scores within the Global Reporting Initiative for a set of global companies. Prediction in Economics has proved to be an involved task. Treating this problem as complex systems is of exceptional interest now when computer technologies provide enormous possibilities for collecting, storing and processing of information obtained by tracing system behavior. The fundamental interest is that results from modeling this problems as complex systems may shed more light on the general aspects of its evolution [12–15]. In this work we

* Corresponding author at: IF, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, RS, Brazil.
E-mail address: david.dominguez@ufrgs.br (D. Dominguez).

http://dx.doi.org/10.1016/j.neucom.2015.05.004
0925-2312/

have introduced a series of alternative approach compared to previous applications of ANN to finance modeling [16–18]. These are namely

- the role of topology on the global GRI interactions,
- the dynamic threshold strategy as function of GRI annually activity reporting,
- the role of offline/online neural update which we suppose is related to the main economic hypothesis: rational/adaptive expectation respectively,
- static and sequential learning which concerns to the agents beliefs about the far/near time period assumptions.

In general, an attractor network is a network of nodes, often recurrently connected, whose time dynamics settle to a stable pattern. That pattern may be stationary, time-varying (e.g. cyclic), or even stochastic-looking (e.g., chaotic). The particular pattern a network settles to is called its "attractor" [19]. A metric ANN is a Hopfield type of attractor network using a metric connectivity distribution, i.e. *small-world*, which has a majority of local links with a moderate number of shortcuts. Recent works [20–26] have shown that metric ANNs are able to sustain the existence of locally organized memories. Thus, metric connectivity is vital for real world implementations, where data patterns are spatially and/or temporally organized [27]. That is, the information is non-uniformly distributed. Metric networks are expected to be more suitable to store and retrieve structured patterns than uniform (either fully connected or random for instance) networks do [28]. It is worth to note that metric networks, being short-range architectures, are much cheaper in terms of the wiring cost than long-range ones. We have shown in Section 2.2 that each geographical region and economical sector have their characteristics range of influence and measure it in terms of the Mutual Information (MI) between enterprises which is in accordance with the observed global behavior [1,4,29].

Our goal is to model and predict the GRI scoring based on the information of this standard since its beginning until 2013 using an attractor network to learn a sequence series. The aspects like the involved metric and correlation matrices (i.e. the network weights for storing sequences) together with the nonlinear dynamics nature of the ANN will be characterized for the prediction of this type of sequential data patterns. The proposed model is described in Section 2, where the neural coding, the network topology, the learning and retrieval dynamics and the information measures are detailed. Then, we present the results for the dynamic attractors for each year of the GRI data in Section 3. In Section 4, we build the network using the mutual information between companies, and present the resulting topology configurations as mesoscopic networks for continents and economic sectors. Finally, in Section 5 we outline the conclusions and discuss the present approach to analyze CSR reporting and the worldwide and economic implications of our findings.

2. Metric ANN model for GRI retrieval

In this section the metric ANN model is rigorously described, starting with the neural coding for the GRI scoring patterns, the topology of the network and its learning (slow) and retrieval (fast) dynamics. Also, the information measures employed to determine the quality of the GRI scoring retrieval is defined.

2.1. Neural coding

The data from the GRI scoring report for each pattern (year) μ is coded to a binary variable $\eta_i^\mu \in \{0,1\}$ meaning non-reporting and reporting respectively, for each enterprise i in the dataset. For instance, a given reporting enterprise time series is translated from $\{\eta_i^\mu\}_{\mu=1}^P = (0,0,G3,C+,3p,\ldots,0)$ to $\vec{\eta} = (0,0,1,1,1\ldots,0)$, for years running from 1999 to 2013, that is for a duration of $P=15$ years. A network with N neurons and a fixed number of $K<N$ synaptic connections per neuron is considered. At any given discrete time t, the network state is defined by the set of N independent binary neurons $\vec{\tau}^t = \{\tau_i^t \in [0,1]; i=0,\ldots,N-1\}$, each one active or inactive denoted respectively by the state 1 or 0. The aim of the network is to retrieve a sequence of patterns (in this case, the consecutive years of sustainability reporting score values) $\{\vec{\eta}^\mu, \mu=1,\ldots,P\}$ that have been stored during a learning process. Each pattern $\vec{\eta}^\mu = \{\eta_i^\mu \in [0,1]; i=1,\ldots,N\}$ is a set of biased binary variables with sparseness probability:

$$p(\eta_i^\mu = 1) = a^\mu, \quad p(\eta_i^\mu = 0) = 1 - a^\mu. \tag{1}$$

For unbiased, i.e. uniform variables, $a^\mu = 1/2$. However for the GRI data, the activity ranges around $a^\mu \sim 0.1$ to $a^\mu \sim 0.6$. The mean activity for each pattern μ is $a^\mu = \sum_i^N \eta_i^\mu / N \equiv \langle \eta^\mu \rangle$. The neural activity for any time t is given by the mean: $q^t = \sum_i^N \tau_i^t / N \equiv \langle \tau^t \rangle$.

2.2. Network topology

The synaptic couplings between the neurons i and j are given by the adjacency matrix

$$J_{ij} \equiv C_{ij} W_{ij}, \quad \mathbf{C} = \{C_{ij} \in [0,1]\}, \tag{2}$$

where the topology matrix $\mathbf{C}$ describes the connection structure of the neural network and $\mathbf{W} = \{W_{ij}\}$ is the matrix of learning weights to be described in Section 2.4. The topology matrix contains two types of links: the local and the random ones, respectively. The local links connect each neuron to its K_l nearest neighbors in a closed ring, while the random links connect each neuron to K_r others uniformly distributed in the network. Hence, the network degree is $K = K_l + K_r$. The network topology is then characterized by two parameters, the *connectivity ratio* γ and the *randomness ratio* ω, which are respectively defined by

$$\gamma = K/N, \quad \omega = K_r/K, \tag{3}$$

where ω plays the role of a rewiring probability in the *small-world* model [30]. The storage cost of this network is $|\mathbf{J}| = N \times K$ if the matrix $\mathbf{J}$ is implemented as an adjacency list, where all neurons have K neighbors. The network topology matrix C_{ij} could also be calculated using the Mutual Information between companies, for continents and economic sectors as explained in Section 4.

2.3. Retrieval dynamics

The task of the network is to retrieve the whole learned sequence of patterns (i.e., the full scoring year sequence for each enterprise) starting from an initial neuron state $\vec{\tau}^0$ which is a given seed year or a state close to it. The retrieval is achieved through the noiseless neuron dynamics:

$$\tau_i^{t+1} = \Theta(h_i^t - \theta_i^t), \tag{4}$$

$$h_i^t \equiv \frac{1}{K} \sum_j J_{ij} \frac{\tau_j^t - q_j^t}{\sqrt{Q_j^t}}, \quad i = 1,\ldots,N, \tag{5}$$

where h_i^t denotes the local field at neuron i and time t, and θ_i is its firing threshold. The local mean neural activity is $q_i^t = \langle \tau^t \rangle_i$, and its variance is $Q_i^t = Var(\tau^t)_i$. The local mean is given by spatial averaging: $\langle f^t \rangle_i \equiv \sum_j C_{ij} f_j^t / N = \sum_{k \in C_i} f_k^t / K$, for any given function f of the neuron sites. Here the step function is used:

$$\Theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{6}$$

For convenience, in this work the normalized variables are used, where the site and time dependence are implicit:

$$\sigma \equiv \frac{\tau - q}{\sqrt{Q}}, \quad q \equiv \langle \tau \rangle, \quad Q \equiv Var(\tau) = q(1-q) \tag{7}$$

$$\xi \equiv \frac{\eta - a}{\sqrt{A}}, \quad a \equiv \langle \eta \rangle, \quad A \equiv Var(\eta) = a(1-a), \tag{8}$$

where a and q are the pattern and neural activities, respectively.

2.4. Learning dynamics

If the network learns a set $P = \alpha K$ of static patterns, $\langle \xi^\mu \xi^\nu \rangle = 0$, these are stored by the network couplings W_{ij} using the classical Hebbian rule [31–33] for the Hopfield model:

$$W_{ij}^{\mu+1} = W_{ij}^{\mu} + \frac{1}{N}\xi_i^\mu \xi_j^\mu \text{ (online)}, \quad W_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \xi_i^\mu \xi_j^\mu \text{ (offline)}. \tag{9}$$

For the case of learning sequential patterns the former Hebbian rule, Eq. (9) is written as

$$W_{ij}^{\mu+1} = W_{ij}^{\mu} + \frac{1}{N}\xi_i^{\mu+1} \xi_j^\mu \text{ (online)}, \quad W_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \xi_i^{\mu+1} \xi_j^\mu \text{ (offline)}. \tag{10}$$

It is worth to note that Eqs. (9) and (10) right parts are considered non-iterative offline learning rules, that is, all patterns must be learned at the beginning of the retrieval process. In contrast, Eqs. (9) and (10) left parts are online iterative rules. The neurons are updated according to Eq. (5), for both static and sequential cases. But, for the sequential rule (related to Eq. (10), online and offline), the retrieval time is the same as the learning step, $t = \mu$.

For the purpose of the GRI reporting analysis done in this work, one may consider the static online dynamics as a feed-forward learning that takes into account only the past events of the sequence. The static offline analysis has no temporal ordering (memoryless). For the case of sequential online dynamics, nearest future expectations will influence the past. The sequential offline dynamics is influenced by the temporal correlations of the sequence but with no preferential temporal ordering (past–future symmetry). The most relevant differences seen in the results in the next section, might be related to the main hypothesis in Economy about agents prediction of future behavior. On the one hand, the rational expectations theory, for which agents are not influenced by time fluctuations, can be modeled by the offline learning rule. On the other hand, the hypothesis of adaptive expectations is a process where agents future predictions are influenced by the near past behavior, similarly as the online learning rule.

2.5. The information measures

In order to evaluate the network retrieval performance, two measures are considered: the global overlap and the load ratio. The overlap is used as a temporal measure of information, which is adequate to describe instantaneously the network ability to retrieve each year of the scoring sequence. In this case, the overlap m_μ^t between the neural state σ^t at time t and the scoring year ξ^μ is

$$m_\mu^t \equiv \frac{1}{N}\sum_i^N \xi_i^\mu \sigma_i^t, \tag{11}$$

which is the normalized statistical correlation between the learned year η_i^μ and the neural state τ_i^t at a given iteration t in the sequence. One lets the network evolve according to Eqs. (4) and (5), and measures the overlap between the network states and the GRI scoring sequence. When the overlap between a given year and the corresponding neural states of the network is $m=1$, the network has retrieved the year without noise. When the global overlap m is zero, the network carries no macroscopic order. In this case, the corresponding year cannot be retrieved. For intermediate values of m, where $0 < m < 1$, the series can be partially recovered with a given level of noise.

For metric connectivity and information structured in blocks (see Tables 1 and 2), it is useful to define the blocks as the structured pieces of information that emerge in the network [22,27]. If the contiguous neurons are distributed within b blocks, one may define mesoscopic parameters restricted to the λth block ($\lambda = 1, \ldots, b$). The *block's overlap* between the neural states and one individual pattern at an unspecified time step is

$$m_\lambda \equiv \frac{1}{L}\sum_{i \in \lambda} \xi_i \sigma_i. \tag{12}$$

One can consider m_λ as a random variable and estimate the average of this variable across the blocks as $\langle f_\lambda \rangle_b \equiv (1/b)\sum_{\lambda=1}^{b} f_\lambda$. Thus, one can define the macroscopic parameters from the mesoscopic measures m_λ. The relevant order parameters measuring the quality of retrieval are the mean (m) and the variance (v) of the block overlap distribution, defined as

$$m \equiv \langle m_\lambda \rangle_b \quad \text{and} \quad v \equiv \langle m_\lambda^2 \rangle_b - m^2. \tag{13}$$

Here, m is the usual global overlap defined in Eq. (11), and the standard deviation $\delta = \sqrt{v}$, which one may call the *block* overlap, conveys the local information for the defined blocks. One may use the normalized form $D = \delta/m$ in order to better depict the yearly evolution of the blocks overlap for the GRI sequence (Figs. 3 and 4).

One is also interested in the load ratio $\alpha \equiv P/K$, that accounts for the storage capacity of the network. When the number of stored patterns increases, the noise due to interference between patterns also increases and the network is not able to retrieve them. Thus, the overlap m goes to zero. A good trade-off between a negligible noise (i.e when $1 - m \sim 0$) and a large scoring sequence (i.e. a high value of α) is desirable for any practical-purpose model.

Table 1
MI topology configuration of companies per geographical region.

Continent	Companies	γ_l	γ_r	ω^s
Africa	370	0.1575	0.0257	0.1402
Asia	1463	0.0679	0.0227	0.2504
Europe	2203	0.0285	0.0372	0.5660
Latin Am.	825	0.0159	0.0368	0.6981
North Am.	772	0.0115	0.0373	0.7644
Oceania	264	0.0183	0.0346	0.6541

Table 2
MI topology configuration of companies per economic sector.

Sector	Companies	γ_l	γ_r	ω^s
Energy	624	0.0694	0.0297	0.3000
Finance	674	0.0634	0.0301	0.3221
Industry	1618	0.0463	0.0293	0.3875
Other	823	0.0123	0.0374	0.7528
Primary	324	0.0187	0.0348	0.6506
Services	1578	0.0123	0.0418	0.7726
Technology	256	0.0117	0.0349	0.7493

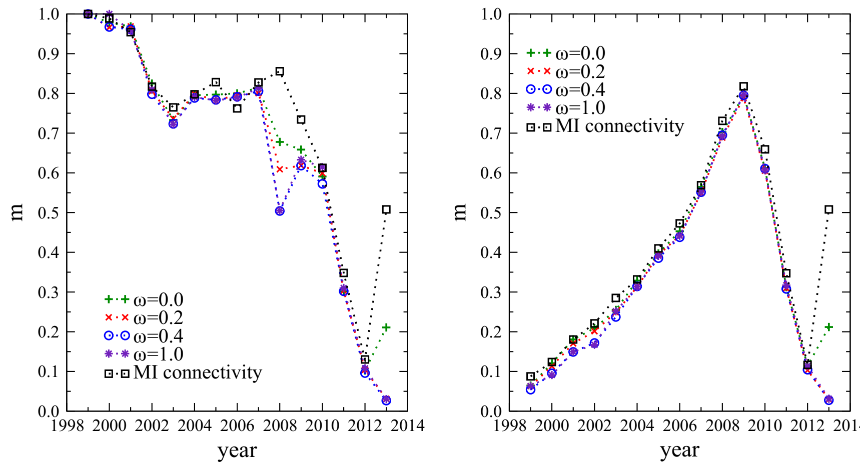

Fig. 1. Global retrieval of GRI scores as static patterns. Network with $N=5897$, $K=200$ and retrieval steps $t=100$ for each pattern (year). Left: Online learning. Right: Offline learning.

3. Results

One has a dataset with $N=5897$ companies and their corresponding CSR reports for the last 15 years, from year 1999 to 2013. The dataset mean sparseness is $a=0.1832$, a threshold value of $\theta=0.83$ was used according to the dynamic threshold model for sparse networks proposed in Dominguez et al. [22]. In order to take advantage of the metric topology of the attractor network, the companies were ordered by regions, in order to capture the inner structured of the data according to their geographical distribution by continents. The network degree is set to $K=200$, for all simulation results. The value of K is bounded from up and down. The upper limit is the minimal block size given by the regions/sectors number of enterprises (see Tables 1 and 2). The lower limit is that for which almost all enterprises remain connected. We have chosen a value closer to the upper bound so that the network is able to learn a large number of patterns, which scales as $P=\alpha K$.

In Fig. 1 is depicted the global retrieval overlap m of static patterns (years), using networks with $N=5897$, $K=200$, for the two extremes of the small-world topology, the regular ring with $\omega=0.0$ and the uniform random network $\omega=1.0$, and two intermediate values $\omega=0.2$ and $\omega=0.4$. The network topology configuration built using the Mutual Information between companies is also depicted (see Section 4). The GRI sequence is stored as static patterns.

In Fig. 1-left, online learning is used to store and retrieve the patterns (years), and the retrieval overlap is depicted. One can appreciate that the network is able to retrieve the series with a good quality ($m>0.5$) until year 2010, when the network saturates and the retrieval quality decays ($m<0.5$). In year 2010, the reporting behavior starts to saturate, the reporting pattern start to repeat, so the retrieval quality worsens due to a high cross talk term noise given by the high correlation of the saturation year (between 2009 and 2010) with the next ones. In Fig. 1-right for offline learning, this peak (2009) is more pronounced. This is because of, for offline learning all GRI reporting years are learned before the retrieval process, and one has the total cross talk term noise value, that is a memoryless process without temporal ordering, not an incremental one as is the case of the online dynamics, which is a feed-forward learning with past memory (Fig. 1-left). One can get valuable information from Fig. 1-right. The first reporting years are very correlated, that is the reporting pattern is more or less the same. However, this rapidly changes as more companies begin to report, and the GRI pattern differentiates from one year to the next one. Once a saturation point in year 2009 is reached, the GRI pattern starts again to repeat and the correlation between years increases, decreasing the retrieval quality (m).

In Fig. 2 is depicted the GRI reporting retrieval overlap m for similar conditions presented in Fig. 1, but this time the network learns the patterns (years) as a sequence. Again, online (left) and offline (right) learning are used. One can appreciate a similar overall behavior as the one depicted in Fig. 1. Notice that the network starts the retrieval process in the exact pattern (beginning year), that is, $m=1$ for year 1999. Also, in Fig. 2-right, the peak is shifted one year (2010), this is due to the fact that the learning rule is now taking into account the interactions between one year and the next, that is years 2009–2010 for the peak. One should notice that the *MI*-topology configuration performs slightly better for the static retrieval than the small-world network (Fig. 1). However, for the temporal sequence, performs even poorly than the regular Ring network of local neighbors $\omega=0.0$ (Fig. 2). Although the sequential learning retrieval (Fig. 2) seems to worsen compared with static learning (Fig. 1), in terms of the attractor dynamics, from a financial point of view, the sequential learning could better represent the GRI reporting behavior before, during, and after the crisis. This is because of the near future expectations are considered for the online dynamics, and for the offline dynamics a temporal correlation albeit symmetric with respect the past and future years.

In Figs. 3 and 4, results are shown for a network with $N=5897$, $K=200$ and *MI*-topology configuration, in order to capture the inner structured of the data according to their geographical distribution (continents) and economic sectors, as detailed in Tables 1 and 2, respectively. The retrieval overlap m is shown for each block (continent or economic sector) for their corresponding GRI reporting years. Again, the behavior in Figs. 3 and 4 is similar to the one shown in Fig. 1. Static patterns are used for online and offline learning, left and right panel respectively. Although the online dynamics seems to perform better (Fig. 3-left), the offline dynamics better represents the behavior around the economic crisis (Fig. 3-right). Here, the normalized deviation D clearly shows the crisis around the years 2007–2009. It is worth to note that in Fig. 4 the Finance sector seems to be gaining momentum in GRI reporting, probably as an effort to gain credibility after the crisis peak.

4. Information and topology

One can find works suggesting that new performance and evaluation tools for Corporate reporting systems are needed [34,35]. In this section we propose to measure the MI between reporting enterprises distributed in blocks, geographically (continents) and economically (sectors), in order to build the network

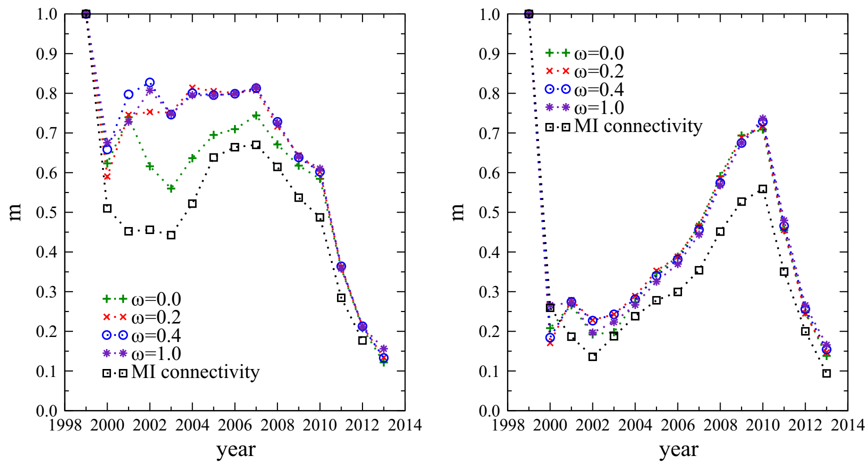

Fig. 2. Global retrieval of GRI scores as sequence of patterns Network with $N = 5897$, $K = 200$, and $cycle = 1$. Left: Online learning. Right: Offline learning.

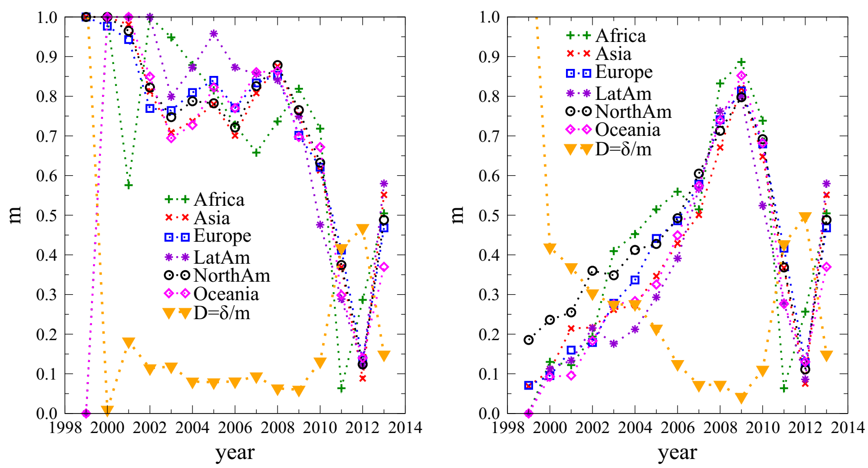

Fig. 3. GRI patterns retrieval by geographic regions. Static patterns retrieval. Network with $N = 5897$, $K = 200$, $t = 100$ with MI connectivity. Left: Online learning. Right: Offline learning.

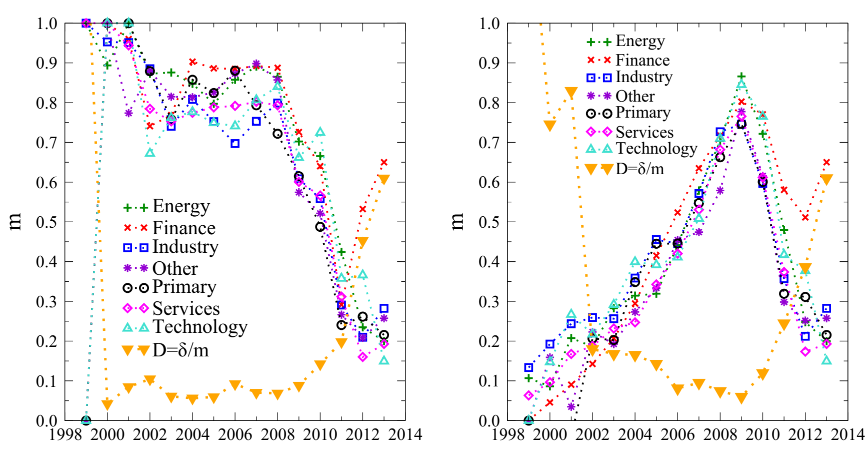

Fig. 4. GRI patterns retrieval by economic sectors. Static patterns retrieval. Network with $N = 5897$, $K = 200$, $t = 100$ with MI connectivity. Left: Online learning. Right: Offline learning.

topology. The MI topology configuration combined with the attractor dynamics of the network, constitute a novel tool for Corporate reporting pattern analysis.

The following calculations will consider the static and offline configuration of the sustainability report. In principle one should compare the present results with the neural network shown in Figs. 1/ 4 (right panels). Nevertheless, the network topology obtained here could also be used for the purpose of the whole previous section: instead of the a priori defined connectivity matrix C_{ij}.

First we briefly define the statistical equations we used here

$$MI[\boldsymbol{\eta}_i;\boldsymbol{\eta}_j] = \langle \ln[p(\boldsymbol{\eta}_i|\boldsymbol{\eta}_j)/p(\boldsymbol{\eta}_i)] \rangle \tag{14}$$

is the mutual information between enterprise given by the vector

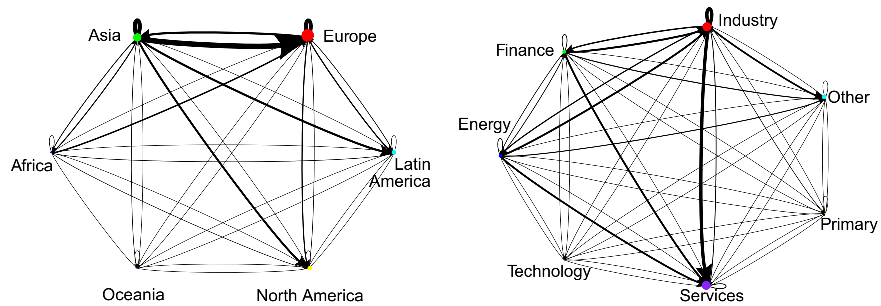

Fig. 5. Network mesoscopic configuration given by the MI between companies. Left: topology configuration for geographic regions. Right: topology configuration for economic sectors.

$\boldsymbol{\eta}_i = \{\boldsymbol{\eta}_i^{\mu}\}_{\mu=1}^{t}$, where t is the time of the observed data, correspond to the years (1999/2013).

The conditional probabilities described above are

$$p(\boldsymbol{\eta}_i|\boldsymbol{\eta}_j) = p(\boldsymbol{\eta}_i, \boldsymbol{\eta}_j)/p(\boldsymbol{\eta}_j), \tag{15}$$

where $p(\boldsymbol{\eta}_i, \boldsymbol{\eta}_j)$ is the joint probability for ith and jth companies to behave similarly with respect to their sustainability performance (i.e. GRI). Here, $p(\boldsymbol{\eta}_i)$ is the marginal probability distribution of a single company.

In order to smoothen the distribution to allow a feasible attractor neural dynamics, instead of using the full 2^P joint distribution configurations [36] for each N^2 couple of enterprises, we estimated it as $p(\boldsymbol{\eta}_i, \boldsymbol{\eta}_j) \propto p(\eta_i^P, \eta_j^P)$ where the latter is the result of the iteration $p(\eta_i^{\mu}, \eta_j^{\mu}) = p(\eta_i^{\mu-1}, \eta_j^{\mu-1}) + 1/P$, for $\mu = 1, \ldots, P$. Here the frequency counts start as $p(\eta_i^0, \eta_j^0) = 0$. For instance, if $P=5$, $\boldsymbol{\eta}_i = (1,0,1,0,1)$ and $\boldsymbol{\eta}_j = (0,1,0,1,1)$, the approximation gives $p(\eta_i^{\mu} = 0, \eta_j^{\mu} = 0) = 0$, $p(\eta_i^{\mu} = 0, \eta_j^{\mu} = 1) = 2/5$, $p(\eta_i^{\mu} = 1, \eta_j^{\mu} = 0) = 2/5$, $p(\eta_i^{\mu} = 1, \eta_j^{\mu} = 1) = 1/5$.

After calculating all the matrix $MI[\boldsymbol{\eta}_i; \boldsymbol{\eta}_j]$, one uses a strategy to build up the connectivity matrix C_{ij}. Among the possibilities one could choose a cutoff M_c, above which one considers that there is a connection between companies ith and jth (below M_c the connection is neglected). Such strategy shows a huge variability from company to company in the number of connection K_i. In order to keep such number constant, one sorted for each company those K neighbors with the highest MI. So, the newly created C_{ij} could be used to: (a) recalculate the neurodynamics shown in Figs. 1/4, (b) describe the topological characteristic of the connectivity C_{ij}. The global parameters of C_{ij} are similar to that defined for an artificial network in Eq. (3), the density and the range of the connections, γ, ω, according to Eq. (16).

However, more important are the local characteristics of C_{ij}: the density and the range of the connections for each sector or region density. The local topology measures one studies are the following:

$$\gamma_r^s = \sum_{i \notin S} K_i^s/[(N-L_s).L_s]; \quad \gamma_l^s = \sum_{i \in S} K_i^s/[(L_s-1).L_s]; \quad \omega^s = \gamma_r^s/(\gamma_r^s + \gamma_l^s). \tag{16}$$

Here S is the group considered (either region or sector); L_s is the size of group S; K_i^s is the connectivity of the companies i with a member inside the group S.

Although the above parameters look slightly different from the usual ω, γ defined for artificial graphs, it is adequate to describe our systems: a large value of the range, $\omega^s \sim 1$, means that the group S is very globalized, while a short range group, $\omega^s \sim 0$, is very localized. In the former case, the information flows more from companies abroad the sector/region, in the latter case, the information remains mostly inside it.

In Tables 1 and 2, we present the parameters ω, γ for each sector/region.

It is worth to note that North America is too global, while Africa is very local. The Energy is a local sector, while Services is a global one. Fig. 5 shows the network mesoscopic topology configuration given by the MI between companies. The mesoscopic blocks (nodes) are presented according to the distribution detailed in Tables 1 (continents) and 2 (economic sectors). It is worth to note that the ω^s parameter, given by the MI topology in Tables 1 and 2, corresponds to the incoming random connection ratio from "distant" nodes (shortcuts from companies in other continents or economic sectors). The size of the nodes is proportional to the number of companies in the corresponding block. The width of the links between nodes is given by companies' connections going from one block to another expressed as a fraction of the total number of connections ($N \times K$). Clearly, the links will be thicker if connecting largest blocks. For instances, in the case of Asia in Fig. 5-left, one can appreciate a thicker self-connection than incoming connections, as expected by the corresponding value of ω^s. Besides, one can appreciate the influences of companies between continents (Fig. 5-left), and between economic sectors (Fig. 5-right), see next section for further discussion on this issue.

5. Discussion and conclusions

Aforementioned findings shed light on diffusion of GRI reporting worldwide and outline an approach to predict its future. First, a global learning is produced. This learning has two patterns (1) among the companies who make the reporting and (2) from the pioneers companies to laggards. In the first case, companies repeat the process year to year. In the second one, companies who make the reporting send a positive signal to the market about the importance to show how sustainable the company is [4]. Thus, pioneer companies could act as a model to other companies. Second, as previous authors suggested, a saturation could be produced. It seems that in the end of 2010 the saturation point is achieved just after the great recession peak, and a new wave was produced. A possible explanation is that more companies are starting to report with this standard and, then, a new learning process could begin. Third, North America and Latin America are strongly influenced by other worldwide regions. Thus, this finding reinforces previous research about the role of Europe as a driver about sustainability and its influence worldwide [4]. Nevertheless, Africa does not receive any influence from other regions or this influence is low. Idemudia [29] mentioned that although a critical CSR research agenda exists with respect to the African context, its advance is slow. Fourth, financial service companies and the energy sector seem to show an independent behavior from other industries. Alonso-Almeida et al. [1] suggested that these sectors could have a specific strategy. Thus, the energy sector could be adopting GRI reporting in an effort to be more sustainable as it is more visible, polluting, and international. On the other hand, the financial sector could seek market credibility and attract new investors after the financial crisis period. Therefore, GRI reporting could help these industries to construct a new identity defined by legitimate behaviors and an improved image. Nevertheless, other industries as services follow a mimetic behavior when identifying the benefits in the market due to efforts to be more sustainable. GRI reporting is continually evolving. For this reason, GRI published its fourth generation of Sustainability Reporting Guidelines, in 2014. GRI believes that G4 will improve sustainability reporting guidance by making it more focused, helping reports to be more relevant. This could mean that changes could be produced in the learning process. A future line of research could analyze this element (GRI G4) for the reporting pattern dynamics.

Acknowledgments

This work has been funded by UNEMI:2014CONVP01-012 and MINECO TIN2010-19607 research grants.

References

[1] María del Mar Alonso-Almeida, Josep Llach, Frederic Marimon, A closer look at the Global Reporting Initiative sustainability reporting as a tool to implement environmental and social policies: a worldwide sector analysis, Corp. Soc. Responsib. Environ. Manag. (2013).
[2] Prado-LorenzoJose-Manuel, Gallego-AlvarezIsabel, Isabel M. Garcia-Sanchez, Stakeholder engagement and corporate social responsibility reporting: the ownership structure effect, Corp. Soc. Responsib. Environ. Manag. 16 (2) (2009) 94–107.
[3] LozanoRodrigo, HuisinghDon, Inter-linking issues and dimensions in sustainability reporting, J. Clean. Prod. 19 (2) (2011) 99–107.
[4] MarimonFrederic, Alonso-AlmeidaMaría del Mar, RodríguezMartha del Pilar, Cortez AlejandroKlender Aimer, The worldwide diffusion of the global reporting initiative: What is the point? J. Clean. Prod. 33 (2012) 132–144.
[5] Stewart Jones, Geoff Frost, Janice Loftus, Sandra van der Laan, An empirical examination of the market returns and financial performance of entities engaged in sustainability reporting, Aust. Account. Rev. 17 (41) (2007) 78–87.
[6] Dror Etzion, Fabrizio Ferraro, The role of analogy in the institutionalization of sustainability reporting, Organ. Sci. 21 (September (5)) (2010) 1092–1107.
[7] A.M. Shahi, B. Issac, J.R. Modapothala, Intelligent corporate sustainability report scoring solution using machine learning approach to text categorization, in: 2012 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT), October 2012, pp. 227–232.
[8] Alan Willis, The role of the global reporting initiative's sustainability reporting guidelines in the social screening of investments, J. Bus. Ethics 43 (2003) 233–237.
[9] Geoff Lamberton, Sustainability accounting: a brief history and conceptual framework, Account. Forum 29 (1) (2005) 7–26.
[10] José M. Moneva, Pablo Archel, Carmen Correa, Gri and the camouflaging of corporate unsustainability, Account. Forum 30 (2) (2006) 121–137.
[11] Stefan Schaltegger, Roger L. Burritt, Sustainability accounting for companies: Catchphrase or decision support for business leaders? J. World Bus. 45 (4) (2010) 375–384 (Sustainable Business).
[12] S. Drożdż, F. Grümmer, A.Z. Górski, F. Ruf, J. Speth, Dynamics of competition between collectivity and noise in the stock market, Physica A: Stat. Mech. Appl. 287 (3–4) (2000) 440–449.
[13] S. Drożdż, F. Grümmer, F. Ruf, J. Speth, Towards identifying the world stock market cross-correlations: Dax versus Dow Jones, Physica A: Stat. Mech. Appl. 294 (1–2) (2001) 226–234.
[14] V.V. Gafiychuk, B.Yo. Datsko, J. Izmaylova, Analysis of data clusters obtained by self-organizing methods, Physica A: Stat. Mech. Appl. 341 (2004) 547–555.
[15] Thomas Bury, Market structure explained by pairwise interactions, Physica A: Stat. Mech. Appl. 392 (6) (2013) 1375–1385.
[16] Emad W. Saad, Danil V. Prokhorov, Donald C. Wunsch, Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks, IEEE Trans. Neural Netw. 9 (6) (1998) 1456–1470.
[17] Kyoung-jae Kim, Artificial neural networks with evolutionary instance selection for financial forecasting, Expert Syst. Appl. 30 (3) (2006) 519–526.
[18] Charles Wong, Massimiliano Versace, Context sensitivity with neural networks in financial decision processes, Global J. Bus. Res. 5 (5) (2011) 27–43.
[19] C. Eliasmith, Attractor network, Scholarpedia 2 (10) (2007) 1380.
[20] David Dominguez, Mario González, Eduardo Serrano, Francisco B. Rodríguez, Structured information in small-world neural networks, Phys. Rev. E 79 (February (2)) (2009) 021909.
[21] Mario González, David Dominguez, Francisco B. Rodríguez, Block attractor in spatially organized neural networks, Neurocomputing 72 (October) (2009) 3795–3801.
[22] David Dominguez, Mario González, Francisco B. Rodríguez, Eduardo Serrano, R. Erichsen Jr., W.K. Theumann, Structured information in sparse-code metric neural networks, Physica A: Stat. Mech. Appl. 391 (3) (2012) 799–808.
[23] K. Koroutchev, E. Koroutcheva, Bump formation in a binary attractor neural network, Phys. Rev. E 73 (2006) 026107.
[24] Yasser Roudi, Alessandro Treves, An associative network with spatially organized connectivity, J. Stat. Mech. 2004 (7) (2004) P07010.
[25] Yasser Roudi, Alessandro Treves, Localized activity profiles and storage capacity of rate-based autoassociative networks, Phys. Rev. E 73 (2006) 061904.
[26] Yasser Roudi, Alessandro Treves, Representing where along with what information in a model of a cortical patch, PLoS Comput. Biol. 4 (3) (2008), e1000012, 03.
[27] Mario González, David Dominguez, Ángel Sánchez, Learning sequences of sparse correlated patterns using small-world attractor neural networks: an application to traffic videos, Neurocomputing 74 (14–15) (2011) 2361–2367.
[28] D. Dominguez, K. Koroutchev, E. Serrano, F.B. Rodríguez, Information and topology in attractor neural networks, Neural Comput. 19 (April) (2007) 956–973.
[29] Uwafiokun Idemudia, Corporate social responsibility and developing countries moving the critical CSR research agenda in Africa forward, Prog. Dev. Stud. 11 (1) (2011) 1–18.
[30] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 422–440.
[31] Daniel J. Amit, Modeling Brain Function: The World of Attractor Neural Networks, Cambridge University Press, New York, NY, USA, 1989.
[32] J. Hertz, J. Krogh, R. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, Boston, 1991.
[33] Thomas Trappenberg, Fundamentals of Computational Neuroscience, Oxford University Press, USA, June 2002.
[34] Francesco Perrini, Antonio Tencati, Sustainability and stakeholder management: the need for new corporate performance evaluation and reporting systems, Bus. Strat. Env. 15 (5) (2006) 296–308.
[35] B.G. Hermann, C. Kroeze, W. Jawjit, Assessing environmental performance by combining life cycle assessment: multi-criteria analysis and environmental performance indicators, J. Clean. Prod. 15 (18) (2007) 1787–1796.
[36] D.R.C. Dominguez, M. Maravall, A. Turiel, J.C. Ciria, N. Parga, Comparison between a replica calculation and an exact solution, Europhys. Lett. 45 (6) (1999) 739.

Mario González received his Ph.D. in Computer Science from the Autonomous University of Madrid (UAM) in 2012. He made a doctoral stay at the Faculty of Engineering of the University of Porto (FEUP), Portugal, funded by the EMECW Lot 20 program. He is currently a lecturer and researcher at Escuela Politécnica Nacional, Ecuador. His research interests cover attractor networks, complex networks and systems.

María del Mar Alonso-Almeida is a Lecturer of Business Administration in the Autonomus University of Madrid, Spain. The author is interested in exploring the corporate governance and sustainability in organizations. The author has written several articles and has participated like coauthor in two books about corporate governance from the strategic point of view and practical. In addition, she is interested in the study of the operations that affect to these companies, as well as in the analysis of their organizational structures and the role of women in management.

Cassio Avila is graduated in physics (2014) and is now a M.Sc. student in the UFRGS, Porto Alegre, RS, Brazil.

David Dominguez received his M.Sc. (1987) and Ph.D. (1993) in physics from the UFRGS, Porto Alegre, Brazil. He worked at KUL, Leuven, Belgium, and at INTA, URJC and UCM, Madrid, Spain. He is a professor at EPS, UAM and IF, UFRGS. His research interests cover statistical mechanics, neural networks, information theory and sociophysics.