INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – May 16-20, 2016

FEATURES OF EDGE-CENTRIC COLLECTIVE DYNAMICS IN MACHINE LEARNING TASKS

Paulo Roberto Urio[1], *Filipe Alves Neto Verri*[2], *Liang Zhao*[3]

[1]Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil, urio@usp.br
[2]Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil, and School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe AZ, USA, filipeneto@usp.br
[3]Ribeirão Preto School of Philosophy, Science and Literature, University of São Paulo, Ribeirão Preto, Brazil, zhao@usp.br

Abstract: We study how collective dynamics can solve Machine Learning tasks. The Edge Domination System is an algorithm to reveal patterns and obtain information of the underlying complex network. The algorithm consists in the simulation of a dynamical system based on particle competition for the dominance of edges. In this paper, we apply this method to semi-supervised tasks of data learning problems. We propose a vertex-centric version of this model and assess the differences between the edge-centric model.

keywords: Analysis and Control of Nonlinear Dynamical Systems with Practical Applications, Complex Networks, Nonlinear Dynamics and Complex Systems, Collective Dynamics and Particle Competition, Machine Learning.

1. INTRODUCTION

We study how collective dynamics can benefit us on data learning tasks. Among the proposed models found in the literature, we have those which feature self-organized and distributed behavior, through the topological structure of the interacting network, and simple evolution rules. Examples of decentralized natural systems include models of ant colonies, birds with flocking behavior, and models of competition for resources and cooperation between species — such models are applied to optimization and learning problems and are increasingly found in the literature [1–3].

A model we have proposed before is the Edge Domination System (EDS) [4], which is a deterministic algorithm used to obtain information of the underlying complex network by exploring patterns revealed by the collective dynamics of particle competition. The model consists of particles randomly walking a complex network. Particles carry fixed labels and try to dominate the edges of the network by passing through more frequently than rival particles. Unlike similar approaches, in this model the resource of competition is the edges instead of vertices. However, the advantages of such approach are not clear. Therefore, in this paper, we explore the differences between resources of competition. We propose a comparable model of particle competition on vertex, and we compare with the EDS.

2. MODEL DESCRIPTION

The Edge Domination System [4] models a particle competition process in a complex network. Particles compete for the domination of edges of the network. The network is a simple, unweighted, undirected graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{L} \cup \mathcal{U}$ is the set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. The sets $\mathcal{L} = \{v_1, \ldots, v_l\}$ and $\mathcal{U} = \{v_{l+1}, \ldots, v_{l+u}\}$ contain the labeled and the unlabeled vertices, respectively. Each labeled vertex $v_i \in \mathcal{L}$ has a label $y_i \in \{1, \ldots, C\}$. The number of labels C is greater than 1. We denote (i, j) as the edge between vertices v_i and v_j. The number of vertices is $|\mathcal{V}| = l + u$, and we suppose that $l \ll u$. The network is represented by the adjacency matrix $A = (a_{ij})$ where $a_{ij} = a_{ji} = 1$ if v_i is connected to v_j.

Particles randomly walk from vertex to vertex while carrying a fixed label. For a particle carrying label c, any particle/vertex of a label different from c is a rival particle/vertex. The edge counts the number of particles that pass through it, separately by each label. The label with the highest count is the label which dominates that edge. In this system, a variable number of particles can be generated into or be removed from the system. The three action rules described below.

Walking. A particle in any vertex chooses the next vertex with equal probability between the neighbors of the current vertex. Let $\deg v_i$ be the degree of v_i. A particle in v_i chooses to move to v_j with probability $a_{ij} / \deg v_i$.

Absorption. The current label domination in that edge determines whether the particle either goes to the vertex or it vanishes from the system (absorbed). This rule applies unless the next vertex is a rival source (sink), in which case the particle is absorbed. Thus, if the destination vertex is not a sink and the rule is applicable, the survival probability is $1-\lambda\sigma_{ij}^c$, where $\lambda \in [0, 1]$ is the competition parameter and σ_{ij}^c is the proportion of rival particles that passed through edge (i, j) in the previous iteration.

Generation. A labeled vertex (source) generates particles according to its degree and the current number of active particles in the system. Let n^c and n_0^c be the current and the initial number of particles carrying label c, respectively, and $\mathcal{G}^c$ be the set of sources for particles carrying label c. If the total number of particles is fewer than the initial number, then each source performs $n_0^c - n^c$ trials with probability $\rho_i^c \propto \deg v_i$ of generating a new particle.

The original model considers an accumulated domination to produce the unfoldings [4]. Conversely, we use only the current domination to simplify our investigation. The convergence of the system justifies this decision because, after several iterations, the current domination becomes proportional to the accumulated domination.

We present now the slightly modified version of the EDS and its vertex-based counterpart model.

2.1. Edge Domination System

The Edge Domination System is a nonlinear Markovian dynamical system with the state and evolution function

$$X(t) = \begin{bmatrix} \boldsymbol{n}^c(t) = \left[n_i^c(t)\right]_{i=1,\dots,|\mathcal{V}|} \\ N^c(t) = \left(n_{ij}^c(t)\right)_{i,j=1,\dots,|\mathcal{V}|} \end{bmatrix}_{c=1,\dots,C} \tag{1}$$

$$\begin{cases} \boldsymbol{n}^c(t+1) = \boldsymbol{n}^c(t) \times P^c\big(N^1(t), \dots, N^C(t)\big) + \boldsymbol{g}^c(\boldsymbol{n}^c(t)) \\ N^c(t+1) = \operatorname{diag} \boldsymbol{n}^c(t) \times P^c\big(N^1(t), \dots, N^C(t)\big), \end{cases} \tag{2}$$

where $P^c(\cdot) = \left(p_{ij}^c(\cdot)\right)_{i,j}$,

$$p_{ij}^c(\cdot) = \begin{cases} 0 & \text{if } v_j \in \mathcal{L} \text{ and } y_j \neq c, \\ \dfrac{a_{ij}}{\deg v_i}\left(1 - \lambda\sigma_{ij}^c(\cdot)\right) & \text{otherwise,} \end{cases} \tag{3}$$

$$\sigma_{ij}^c(N^1, \dots, N^C) = \begin{cases} 1 - \dfrac{n_{ij}^c + n_{ji}^c}{\sum_{q=1}^C n_{ij}^q + n_{ji}^q} & \text{if } \sum_{q=1}^C n_{ij}^q + n_{ji}^q > 0, \\ 1 - \dfrac{1}{C} & \text{otherwise,} \end{cases} \tag{4}$$

and $\boldsymbol{g}^c(\boldsymbol{n}) = \left[g_i^c(\boldsymbol{n})\right]_i$, where

$$g_i^c(\boldsymbol{n}) = \rho_i^c \max\{0,\ 1 - \boldsymbol{1} \cdot \boldsymbol{n}\}. \tag{5}$$

The initial state of the system X is given by an arbitrary discrete *distribution* $\boldsymbol{n}^c(0)$ of initial active particles and $n_{ij}^c(0) = 0$ for all i, j, c.

We group the edges by label domination at the current time. For each label c, the subset of edges $\mathcal{E}^c(t) \subseteq \mathcal{E}$ is

$$\mathcal{E}^c(t) = \left\{ ij \,\middle|\, \arg\max_q \big(n_{ij}^q(t) + n_{ji}^q(t)\big) = c \right\}. \tag{6}$$

Thus, the unfolding G^c that comprises the edges dominated by the label c at time t is $G^c(t) = (\mathcal{V}, \mathcal{E}^c(t))$.

2.2. Vertex Domination System

Based on the EDS, we model a vertex-centric system with state and evolution function

$$X(t) = \left[\boldsymbol{n}^c(t) = \left[n_i^c(t)\right]_{i=1,\dots,|\mathcal{V}|}\right]_{c=1,\dots,C} \tag{7}$$

$$\boldsymbol{n}^c(t+1) = \boldsymbol{n}^c(t) \times P^c(\boldsymbol{n}^1(t), \dots, \boldsymbol{n}^C(t)) + \boldsymbol{g}^c(\boldsymbol{n}^c(t)), \tag{8}$$

where $P^c(\cdot)$ is eq. (3) with

$$\sigma_{ij}^c(\boldsymbol{n}^1, \dots, \boldsymbol{n}^C) = \begin{cases} 1 - \dfrac{n_j^c}{\sum_{q=1}^C n_j^q} & \text{if } \sum_{q=1}^C n_j^q > 0, \\ 1 - \dfrac{1}{C} & \text{otherwise,} \end{cases} \tag{9}$$

and $\boldsymbol{g}^c(\boldsymbol{n})$ is eq. (5).

The initial state of the system X is given by an arbitrary discrete *distribution* $\boldsymbol{n}^c(0)$ of initial active particles.

3. COMPUTER SIMULATIONS

In this section, we present empirical experiments that address the differences between both systems — Edge Domination System (EDS) and Vertex Domination System (VDS).

First, we formulate a simple example where the system behaves quite differently. Afterward, we compare the performance of each system in artificial semi-supervised tasks. Finally, we measure properties of both systems over time.

3.1. Motivation

Consider the network whose vertices are points in a 2-dimensional plane depicted in fig. 1a. The vertices of this network can be either positive (green, light, shaping a vertical pattern) or negative (red, dark, shaping a horizontal pattern). The central vertex (blue, dark) does not have a defined label. Four vertices (square-shaped) are labeled. We connect each vertex to its nearest neighbor. Using this network as the input of EDS and VDS, one shall expect the fraction of positive and negative particles to be the same in the central vertex for both systems. Figure 2a demonstrates this behavior.

Now, consider adding a new vertex to this dataset as shown in fig. 1b. Notice that the modification barely changes

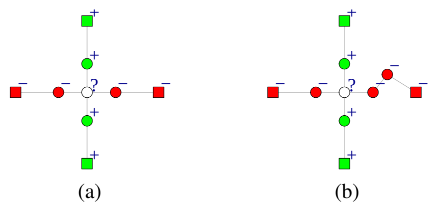

Figure 1 – Toy networks with (a) 9 vertices and (b) 10 vertices. Square-shaped vertices are initially labeled.

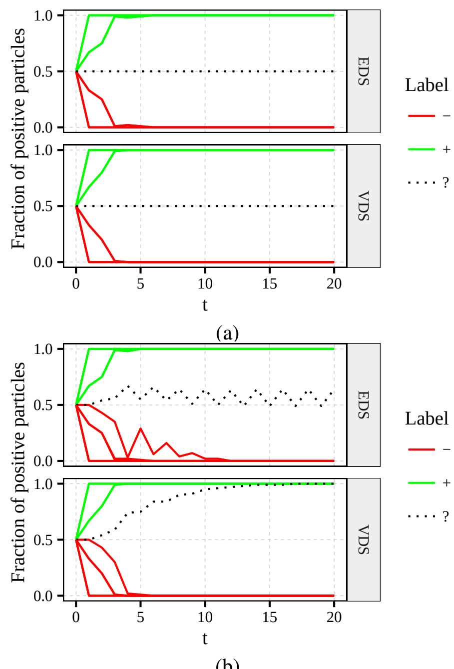

Figure 2 – Fraction of positive particles in each vertex over time with (a) 9-vertices and (b) 10-vertices networks as input.

the network's topology. However, using this network as input both systems behaves quite differently, see fig. 2b. On the one hand, the proportion of different kinds of particles in the central vertex alternates about the same in the edge-centric system. On the other hand, the central vertex is dominated by positive-labeled particles in the vertex-centric system.

This example motivates us to investigate the advantages of the edge-centric model against the vertex-centric one.

3.2. Systems' comparison

Before comparing both systems, we propose a random network with labeled vertices.

The undirected random network $G(n, m, p, q)$ has $n + m$ vertices of which n are positive and m negative labeled. Let $y_i \in \{+, -\}$ be the label of vertex v_i, the probability of existing an edge between (i, j) is p if $y_i = y_j$, and q if $y_i \neq y_j$. Networks G that $n, m > 0$, $p, q \in [0, 1]$, and $p > q$ are a good representation of common machine learning tasks.

We compare classification accuracy of the systems for varying levels of difficulty. The mean accuracy is calculated over 30 independently sampled networks $G(250, 250, 0.6, q)$ for all $q \in \{0.1, 0.2, \ldots, 0.6\}$. We fixed $\lambda = 1$ and set the fraction of initially labeled vertices in $\{0.01, 0.02, 0.05\}$.

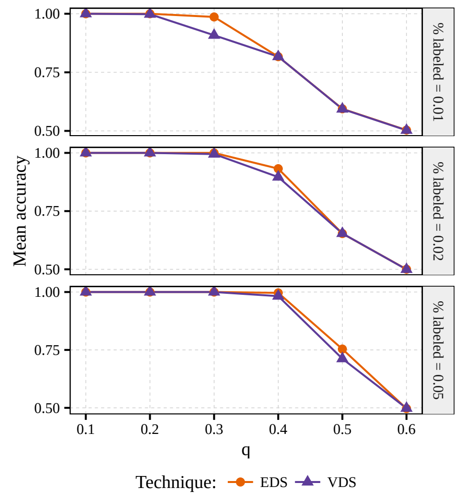

Figure 3 – Mean accuracy of both systems in different networks varying the fraction of labeled input vertices.

To classify an unlabeled vertex v_i, we adopt the original strategy for the EDS at time $\tau = 500$, $y_i = \arg\max_c |\mathcal{E}(\mathcal{N}_i^c(\tau))|$, where $|\mathcal{E}(\mathcal{N}_i^c(\tau))|$ stands for the number of edges in the neighborhood of v_i in $G^c(\tau)$.

In VDS, we just consider the dominating label of the vertex at time $\tau = 500$, $y_i = \arg\max_c n_i^c(\tau)$.

The accuracy results varying the network G and the number of sources is shown in fig. 3. Consistently, as q increases and the number of sources decreases, the problem becomes harder, and the systems obtain lower accuracies.

When $p = q = 0.6$, both systems fail to discriminate the label of the vertices because the network has no meaningful topology. Besides the number of sources compensating the difficulty of the problem, EDS performs better in realistic scenarios where sources are enough to discover a non-trivial class topology.

Now, we detail some general dynamical properties of the systems. Let $\delta(\cdot)$ be 1 for a true argument and 0 otherwise; the properties are listed below.

Positive domination. Defined as the fraction of vertices with the number of positive particles greater than the negative ones,

$$\frac{1}{|\mathcal{V}|} \sum_i \delta\big(n_i^+(t) > n_i^-(t)\big) . \tag{10}$$

Active particles. Defined as the sum of active particles in the system,

$$\sum_i n_i^+(t) + n_i^-(t) . \tag{11}$$

Mean entropy. Defined as the average of the entropy considering the labels of active particles in each vertex,

$$-\frac{1}{|\mathcal{V}|} \sum_i n_i^+(t) \log_2 n_i^+(t) + n_i^-(t) \log_2 n_i^-(t) . \tag{12}$$

The experimental results for the three properties over time are shown in fig. 4. Each result corresponds to the average of 30 independent simulations with $\lambda = 1$ and networks

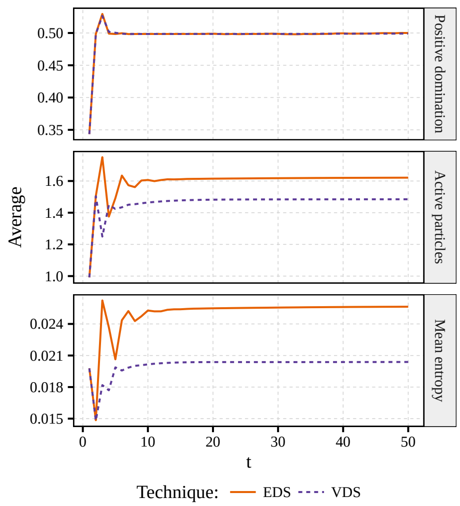

Figure 4 – Properties of EDS and VDS over time.

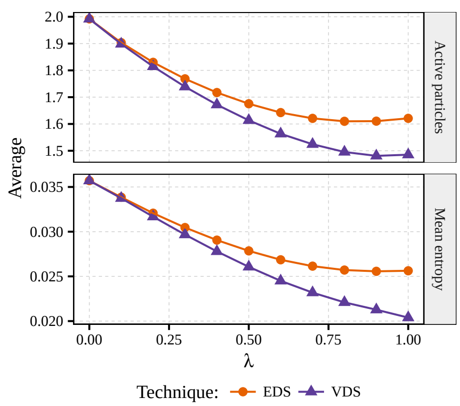

Figure 5 – Active particles and mean entropy properties varying the survival parameter λ.

$G(250, 250, 0.6, 0.3)$ with one positive and one negative labeled vertex. Since the network has the same size and topology for the two labels, half network is dominated by each label in both systems. However, the number of active particles and the entropy are higher in the edge-centric system.

High entropy suggests that particles propagate farther, reaching larger regions dominated by rival particles. As a result, a greater exploration behavior emerges, and particles are more likely to learn from the network's topology. Conversely, a large number of active particles means that fewer particles are absorbed — it could be a consequence of better boundaries between positive and negative labeled vertices, justifying the better accuracy of EDS in these networks.

Using the same network configuration and the number of executions, we also study how two properties — active particles and mean entropy — behave according to the setting of the survival parameter λ at time $t = 50$. The results are shown in fig. 5. On the one hand, as $\lambda \to 0$, the two properties in EDS and VDS become identical, which is expected since the two systems operate identically, without competition. On the other hand, as λ approaches 1, EDS presents a higher number of active particles and higher entropy.

4. FUTURE WORKS

In the previous section, we presented experimental investigations that address the differences between the systems EDS and VDS. In future works, we will provide analytical studies that explain the advantages of EDS over VDS. To do so, we will consider two simplifications in the systems: (a) ignore sink absorption, and (b) suppose that $n_{ij}^c(t) + n_{ji}^c(t) > 0$ for all i, j, t for at least one label c.

Since we expect few labeled vertices, we can neglect the sinks effect in the system. The second simplification means that we expected that there are no regions in the network that lack particles walking in. The resulting system evolution equations for EDS and VDS, respectively, are

$$\begin{cases} n_i^c(t+1) = \rho_i^c \left(1 - \sum_j n_j^c(t)\right) + \sum_j n_{ji}^c(t+1) \\ n_{ij}^c(t+1) = n_i^c(t)\frac{a_{ij}}{\deg v_i}\left(1 - \lambda\left[1 - \frac{n_{ij}^c(t) + n_{ji}^c(t)}{\sum_q n_{ij}^q(t) + n_{ji}^q(t)}\right]\right) \end{cases}, \quad (13)$$

$$\sum_j n_j^c(t)\frac{a_{ji}}{\deg v_j}\left(1 - \lambda\left[1 - \frac{n_i^c(t)}{\sum_q n_i^q(t)}\right]\right). \quad (14)$$

5. CONCLUSION

In this paper, we present a vertex version of the model of particle competition, which is similar to the already proposed Edge Domination System. We also perform experiments to study their differences.

Since both systems have the same time and memory complexity order, we conclude that using EDS is preferred over VDS, as the former have better features for learning tasks.

ACKNOWLEDGMENTS

This research was supported by the São Paulo State Research Foundation (FAPESP), the Coordination for the Improvement of Higher Education Personnel (CAPES), and the Brazilian National Research Council (CNPq).

References

[1] R. A. Gueleri, T. H. Cupertino, A. C. de Carvalho, and L. Zhao, "A flocking-like technique to perform semi-supervised learning," in *IJCNN*, pp. 1579–1586, 2014.

[2] Z. Zhang, K. Long, J. Wang, and F. Dressler, "On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 513–537, 2014.

[3] F. Breve, M. G. Quiles, and L. Zhao, "Interactive image segmentation using particle competition and cooperation," in *IJCNN*, pp. 1–8, IEEE, 2015.

[4] F. A. N. Verri, P. R. Urio, and L. Zhao, "Network Unfolding Map by Edge Dynamics Modeling," 2016. arXiv:1603.01182.