2012 IEEE World Congress on Computational Intelligence

Particle Competition and Cooperation in Networks for Semi-Supervised Learning with Concept Drift

Fabricio Breve[1,2] fabricio@icmc.usp.br

Liang Zhao[2] zhao@icmc.usp.br

[1] Department of Statistics, Applied Mathematics and Computation (DEMAC), Institute of Geosciences and Exact Sciences (IGCE), São Paulo State University (UNESP), Rio Claro, SP, Brazil

[2] Department of Computer Science, Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, SP, Brazil

Outline

- Motivation
- Proposed Method
- Computer Simulations
- Conclusions

Motivation

- Data sets under analysis are no longer only static databases, but also data streams in which concepts and data distributions may not be stable over time.
  - Examples:
    - Climate Prediction
    - Fraud Detection
    - Energy Demand
    - Many other real-world applications

Motivation

- Concept Drift
  - Nonstationary learning problem over time.
  - Learning algorithms have to handle conflicting objectives:
    - Retain previously learned knowledge that is still relevant.
    - Replace any obsolete knowledge with current information.
  - However, most learning algorithms produced so far are based on the assumption that data comes from a fixed distribution.

[1] I. Zliobaite, "Learning under concept drift: an overview," *CoRR*, vol. abs/1010.4784, 2010. **[2]** A. Tsymbal, M. Pechenizkiy, P. Cunningham, and S. Puuronen, "Dynamic integration of classifiers for handling concept drift," *Inf. Fusion*, vol. 9, pp. 56–68, January 2008. **[3]** G. Ditzler and R. Polikar, "Semi-supervised learning in nonstationary environments," in *Neural Networks (IJCNN), The 2011 International Joint Conference on*, 31 2011-aug. 5 2011, pp. 2741 –2748. **[4]** L. I. Kuncheva, "Classifier ensembles for detecting concept change in streaming data: Overview and perspectives," in *Proc. 2nd Workshop SUEMA 2008 (ECAI 2008)*, Patras, Greece, 2008, pp. 5–10. **[5]** A. Bondu and M. Boull´e, "A supervised approach for change detection in data streams," in *Neural Networks (IJCNN), The 2011 International Joint Conference on*, 31 2011-aug. 5 2011, pp. 519 – 526.

Motivation

- Why Semi-Supervised Learning to handle concept drift?
  - Some concept drifts applications requires fast response, which means an algorithm must always be (re)trained with the latest available data.
  - Process of labeling data is usually expensive and/or time consuming when compared to unlabeled data acquisition, thus only a small fraction of the incoming data may be effectively labeled.

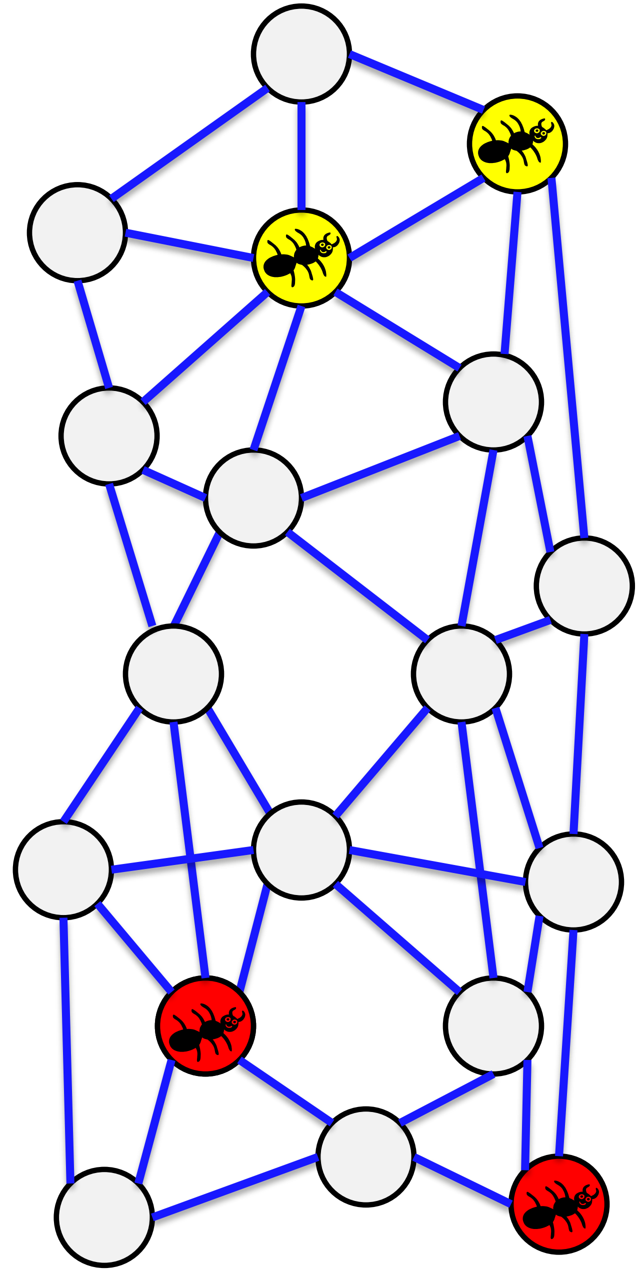
[17] X. Zhu, “Semi-supervised learning literature survey,” Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005. **[18]** O. Chapelle, B. Sch¨olkopf, and A. Zien, Eds., *Semi-Supervised Learning*, ser. Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press, 2006. **[19]** S. Abney, *Semisupervised Learning for Computational Linguistics*. CRC Press, 2008.

Proposed Method

- Particles competition and cooperation in networks.
  - Cooperation among particles representing the same team (label / class).
  - Competition for possession of nodes of the network.
- Each team of particles…
  - Tries to dominate as many nodes as possible in a cooperative way.
  - Prevents intrusion of particles from other teams.

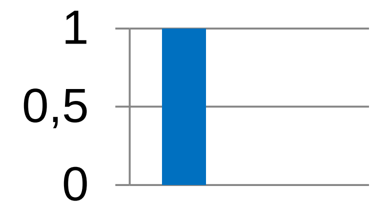
Initial Configuration

- Each data item is transformed into an undirected network node and connected to its *k*-nearest neighbors.
- A particle is generated for each labeled node of the network.
- Particles with same label play for the same team.
- When network maximum size is reached, older nodes are labeled and removed as new nodes are created.
- When maximum amount of particles is reached, older particles are removed as new particles are created.

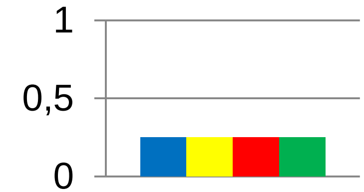


Initial Configuration

- Particles initial position are set to their corresponding nodes.
- Nodes have a domination vector.
  - Labeled nodes have ownership set to their respective teams.
  - Unlabeled nodes have levels set equally for each team.



Ex: [1.00 0.00 0.00 0.00]
(4 classes, node labeled as class A)

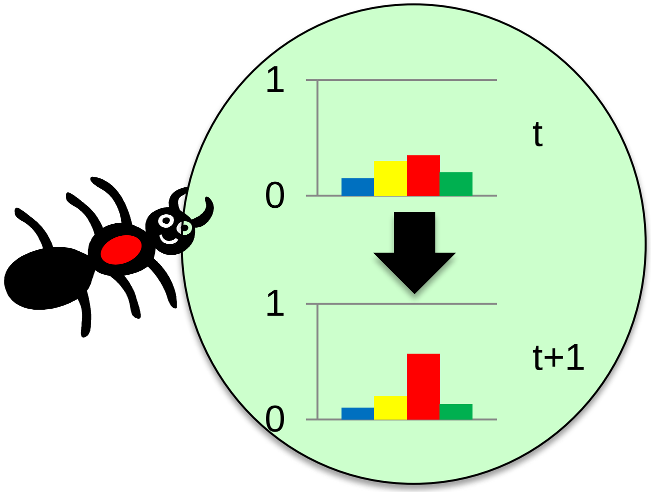


Ex: [0.25 0.25 0.25 0.25]
(4 classes, unlabeled node)

$$v_i^{\omega_\ell}(t) = \begin{cases} \frac{1}{c} & \text{if } y_i = 0 \\ 1 & \text{if } y_i = \ell \\ 0 & \text{otherwise} \end{cases}$$

Node Dynamics

- When a particle selects a neighbor node to visit:
  - It decreases the domination level of the other teams.
  - It increases the domination level of its own team.

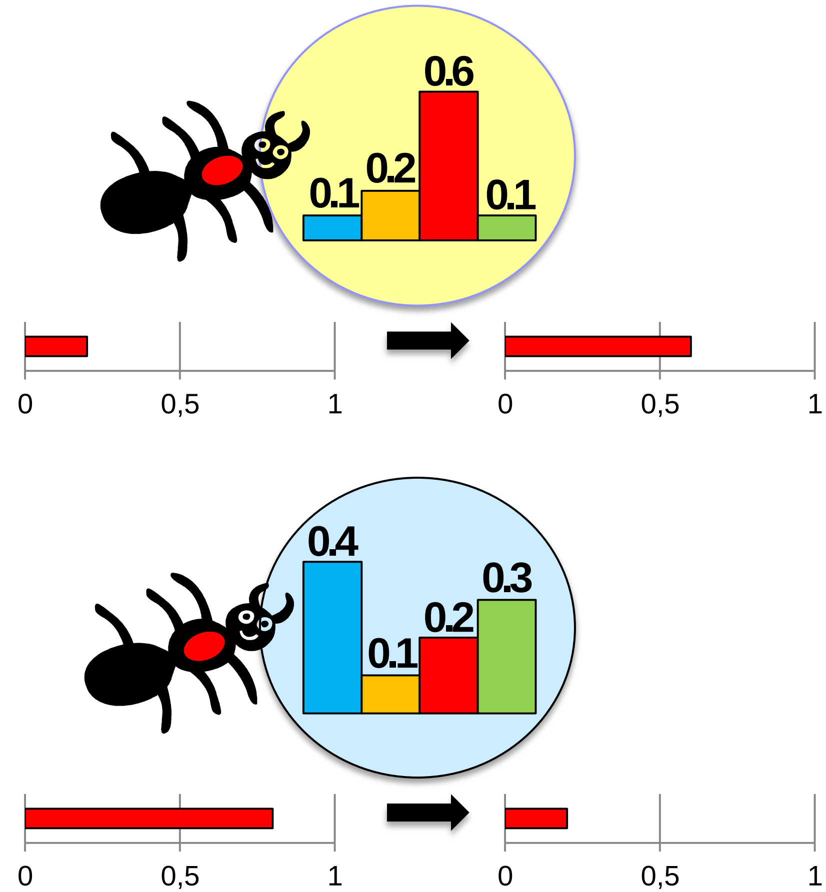


$$v_i^{\omega_\ell}(t+1) = \begin{cases} \max\{0, v_i^{\omega_\ell}(t) - \frac{\Delta_v \rho_j^{\omega}(t)}{c-1}\} \\ \quad \text{if } y_i = 0 \text{ and } \ell \neq \rho_j^f \\ v_i^{\omega_\ell}(t) + \sum_{q \neq \ell} v_i^{\omega_q}(t) - v_i^{\omega_q}(t+1) \\ \quad \text{if } y_i = 0 \text{ and } \ell = \rho_j^f \\ v_i^{\omega_\ell}(t) \quad \text{if } y_i \neq 0 \end{cases}$$

Particle Dynamics

- A particle gets:
  - stronger when it selects a node being dominated by its team.
  - weaker when it selects node dominated by other teams.

$$\boxed{\rho_j^{\omega}(t+1) = v_i^{\omega_\ell}(t+1)}$$

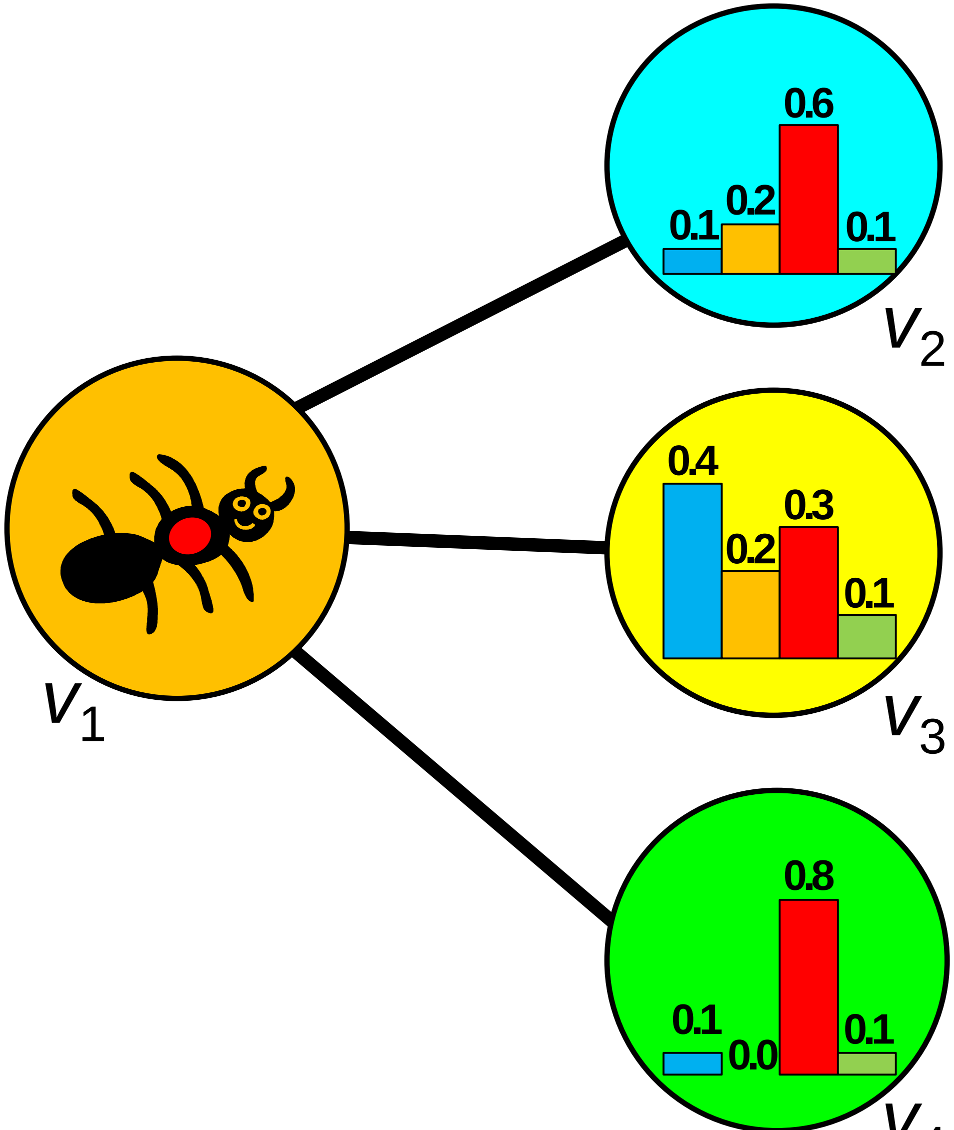


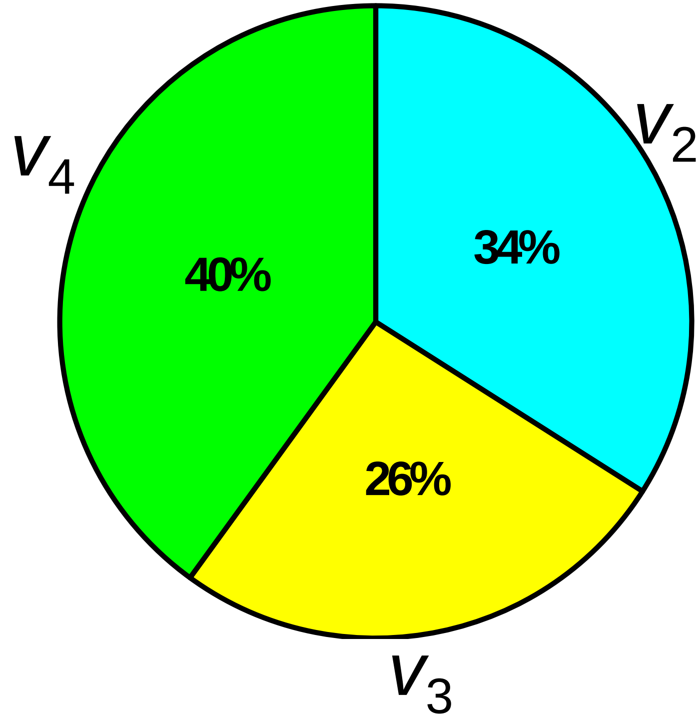
Particles Walk

- Random-greedy walk
  - The particle will prefer visiting nodes that its team already dominates.

$$\boxed{p(v_i|\rho_j) = (1-\alpha)\frac{W_{qi}}{\sum_{\mu=1}^{n} W_{q\mu}} + \alpha\frac{W_{qi}v_i^{\omega_\ell}}{\sum_{\mu=1}^{n} W_{q\mu}v_i^{\omega_\ell}}}$$

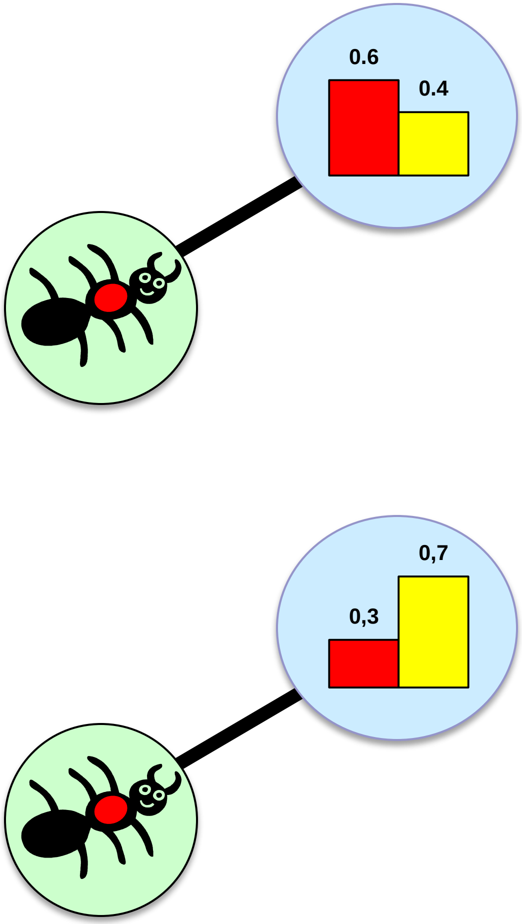
Moving Probabilities



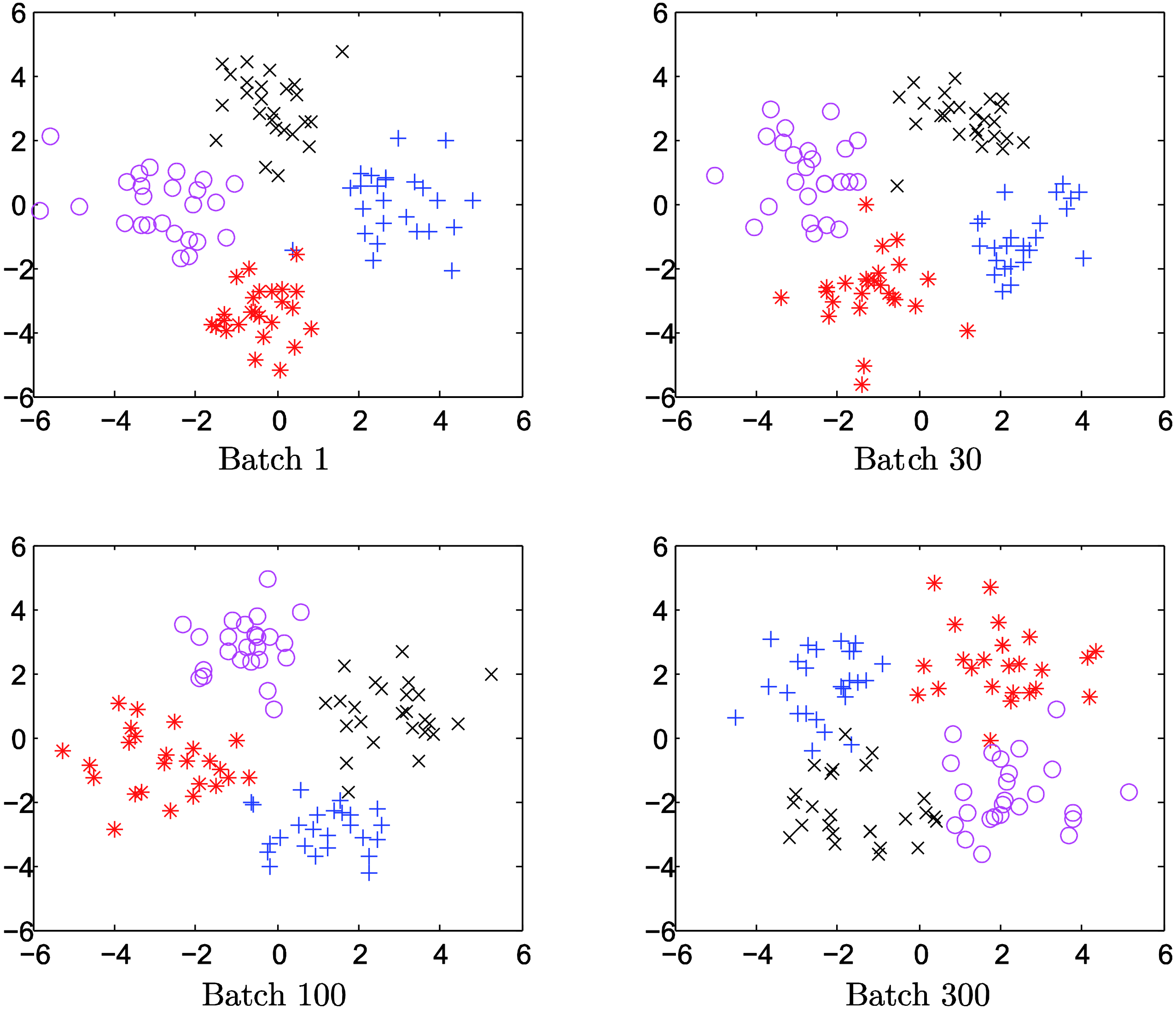


Particles Walk

- Shocks
  - A particle really visits the selected node only if the domination level of its team is higher than others;
  - Otherwise, a shock happens and the particle stays at the current node until next iteration.

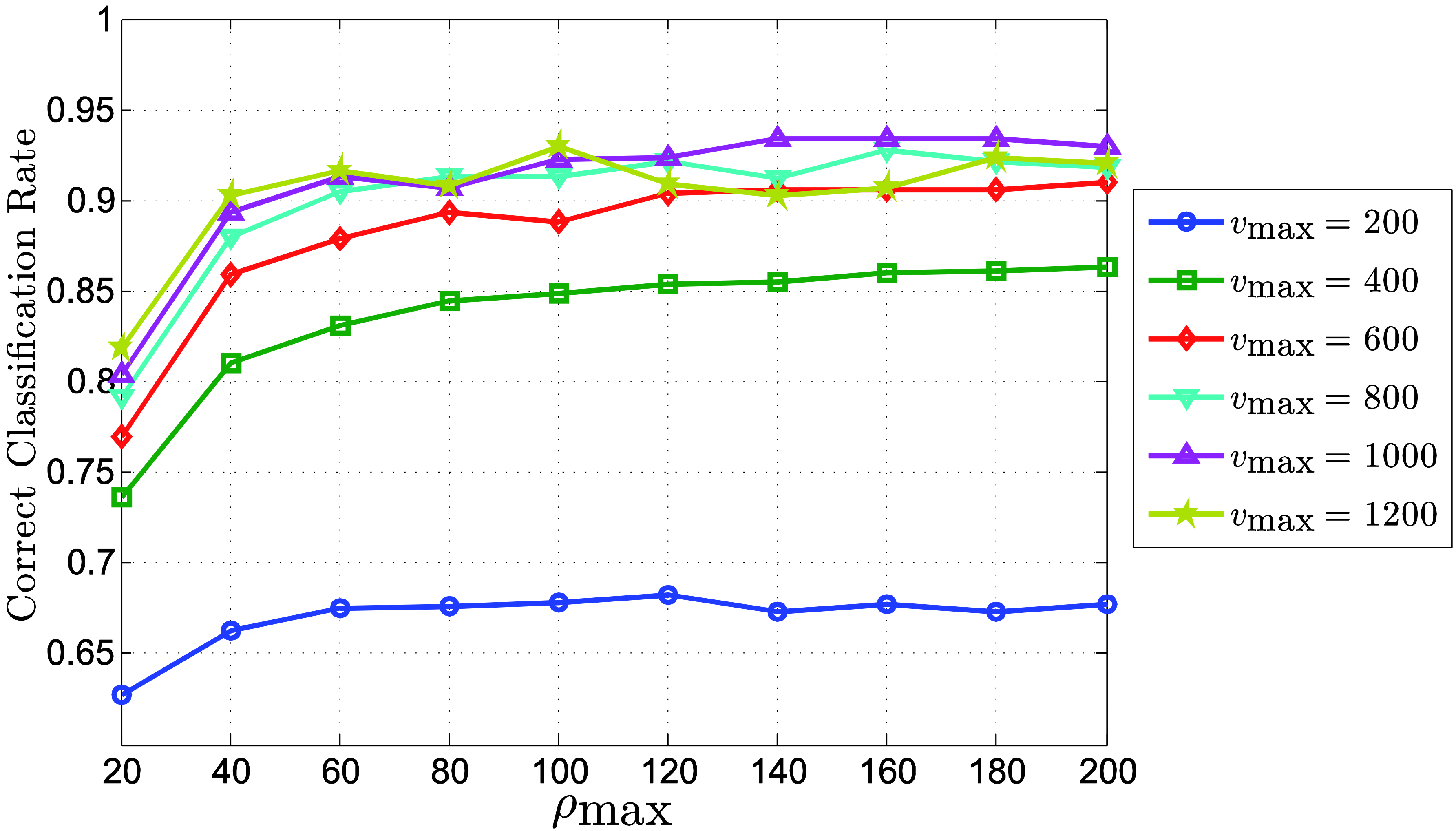


Computer Simulation 1 – Slow Concept Drift



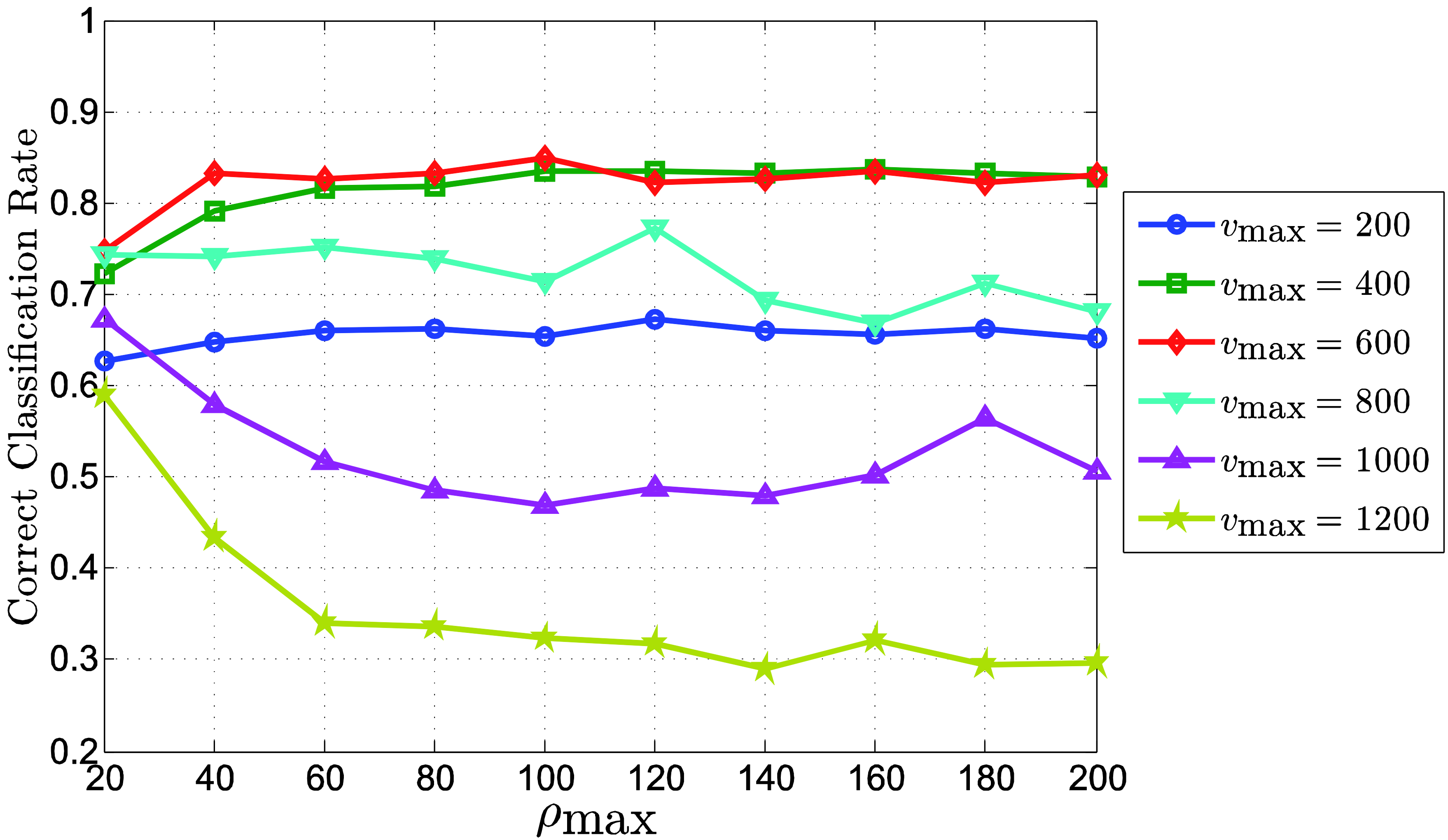
- 50,000 data items.
- 500 batches.
- 100 data items in each batch.
- Data items generated around 4 Gaussian kernels moving clockwise.
- 100,000 particle movements between each batch arrival.
- 10% labeled data items, 90% unlabeled.
- *k* = 5.

Computer Simulation 1 – Slow Concept Drift



Simulation 1: Slow Concept Drift. Correct Classification Rate with varying maximum network size (v_{max}) and maximum amount of particles (ρ_{max}). ***n* = 50,000**.

Computer Simulation 2 – Fast Concept Drift



Simulation 2: Fast Concept Drift. Correct Classification Rate with varying maximum network size (v_{max}) and maximum amount of particles (ρ_{max}). ***n* = 10,000**.

Conclusions

- New biologically inspired method for semi-supervised classification in nonstationary environments.
  - Specially suited for gradual or incremental changes in concept.
  - Passive concept drift algorithm.
    - Naturally adapts to changes.
    - No explicit drift detection mechanism.
  - Does not rely on base classifiers with explicit retraining process.
    - Built-in mechanisms provide a natural way of learning from new data, gradually “forgetting” older knowledge.
  - Single classifier approach.
    - Most other passive methods rely on classifier ensembles.

Future Work

- Build mechanisms to automatically select the parameters which control the sizes of the network and the set of particles, according to the data that is being fed to the algorithm.
  - This could highly improve the performance of the algorithm in scenarios where the concepts may be stable for sometime and/or have different drift speeds through time.

Acknowledgements

- This work was supported by:
  - State of São Paulo Research Foundation (FAPESP)
  - Brazilian National Council of Technological and Scientific Development (CNPq)
  - Foundation for the Development of Unesp (Fundunesp)

FAPESP

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

Fundunesp
Fundação para o Desenvolvimento da UNESP

2012 IEEE World Congress on Computational Intelligence

Particle Competition and Cooperation in Networks for Semi-Supervised Learning with Concept Drift

Fabricio Breve[1,2] fabricio@icmc.usp.br
Liang Zhao[2] zhao@icmc.usp.br

[1] Department of Statistics, Applied Mathematics and Computation (DEMAC), Institute of Geosciences and Exact Sciences (IGCE), São Paulo State University (UNESP), Rio Claro, SP, Brazil

[2] Department of Computer Science, Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, SP, Brazil