A supervised learning approach based on STDP and polychronization in spiking neuron networks

Hélène Paugam-Moisy¹, Régis Martinez¹ and Samy Bengio²

¹LIRIS - CNRS - Université Lumière Lyon 2
Lyon, France
http://liris.cnrs.fr

²IDIAP Research Institute
Martigny, Switzerland
http://www.idiap.ch
Samy is now at Google
Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion
Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion
Motivation

- In **Spiking Neuron Networks** (SNNs), information processing is based on the times of spike emissions.

- SNNs are a very powerful new generation of artificial neural networks but efficient learning in SNNs is not straightforward.

- A current track is to simulate the synaptic plasticity, as can be observed by neurobiologists [Bi and Poo, 1998] but this method lacks supervised control of learning.
Theoretical fundations

- Theoretically, the use of **delays** increases the learning capacity of SNNs...
  [Maass, 1997] [Schmitt, 1999]

  ... but delays are rarely used in SNN models

- Recent advances in neural networks (ESN [Jaeger, 2001], LSM [Maass et al, 2002]) give interesting results

- The concept of **polychronization** emphasizes the importance of delays for explaining neural activity
  [Izhikevich, 2006]
Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion

Régis Martinez

A supervised learning approach based on STDP and polychronization in spiking neuron networks
A better computational power is a good point, but what about the learning algorithm? How to take advantage of the computational power of delays?

- We take advantage of polychronous groups activations to monitor activity in the network
- We define a supervised\(^1\) learning mechanism to control the computational power of a SNN

Polychronization will help us monitor and understand the network activity.

\(^1\)simplest way for us to show that polychronization can actually be a reliable information coding
Motivations
Problematics
Network architecture
Learning mechanisms
Results (1)
Polychronization
Results (2)
Conclusion

Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion
The model

- Maintains biological plausibility within the internal network
- Neuron model: Spike Response Model ($SRM_0$) [Gerstner 1997]
- Inspired from LSM/ESN architectures:
  - input layer of spiking neurons
  - recurrent randomly connected internal network
  - output layer which supports a supervised learning rule

```
K input cells

Internal network:
- M internal cells
- input connections
- internal connections
- output connections with adaptable delay

2 output cells
```

```
class 2
class 1
```
The model

Input layer (stimulation layer):
- 10 neurons
- Input injection

Internal network

K input cells

M internal cells

2 output cells

input connections

internal connections

output connections with adaptable delay
The model

K input cells

Internal network

M internal cells

2 output cells

- input connections
- internal connections
- output connections with adaptable delay

Internal Network:

- 100 neurons, 80% excitatory, 20% inhibitory
- Random recurrent topology
- Connection delays fixed (but randomly chosen) between 1 and 20 ms
The model

K input cells

M internal cells

2 output cells

Input connections

Internal connections

Output connections with adaptable delay

Output layer:

- 2 neurons: one for each target class
- Receives a connection from each internal neuron
The model

Tested on a classification task

Two input patterns:
Target neuron must fire before non-target neuron
Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion
A two scale learning algorithm

1. **Unsupervised** learning: Spike Time Dependent Plasticity (STDP) within the internal network (ms time scale) [Kempter et al., 1999]

2. **Supervised** mechanism: delay adaptation on output connections (at each input presentation) based on a margin criterion [Vapnik, 95]
1. Unsupervised learning algorithm

**Unsupervised** learning: Spike Time Dependent Plasticity (STDP) within the internal network (ms time scale)

- Temporal hebbian rule, suitable for SNNs
- At the synaptic level (local mechanism)
- Depending on activity going through the synapse
- Causality based on spike emissions order
2. Supervised learning algorithm

After the presentation of a given input pattern $p$, if target/non-target spikes order is OK AND if margin between target/non-target spikes $> \epsilon$ then pattern is well classified.

Otherwise,
- for target neuron: decrement the delay ($-1 ms$)
- for non-target neuron: increment the delay ($+1 ms$)
Régis Martinez

A supervised learning approach based on STDP and polychronization in spiking neuron networks
Simulation protocol

- Initial noisy stimulation: noise presented during 300 ms
- Learning phase: alternated presentation of two patterns
- Generalization phase: alternated presentation of the two noisy patterns

NB: One presentation every 100 ms
Initialization phase

![Graph showing neuron activity over time.](image-url)
Learning phase observation

- Decreasing internal activity (STDP)
- Activity pattern different from an input to the other
- Margin evolution
Generalization performance

- Error rate with noise 4: 4%
- Error rate with noise 8: 19%
- Hard to discriminate by human
Motivations

Problematics

Network architecture

Learning mechanisms

Results (1)

Polychronization

Results (2)

Conclusion
**Polychronization [Izhikevich, 2006]**

**Definition**: neuron interactions characterized by spike times following a precise temporal pattern, depending on delays.

**Example**:

If \( N_1 \) emits a spike at \( t \), and \( N_3 \) at \( t + 7 \), then \( N_2 \) emits a spike at \( t + 15 \).

A set of such interacting neurons is called a *polychronous group*.
Scanning for supported polychronous groups

Structure
Polychronous groups are supported by the topology.
- connections between neurons
- delays of the connections
- A given topology = a particular set of supported polychronous groups
- Each neuron can be involved in several polychronous groups

To find all supported polychronous groups, we use the same algorithm as [Izhikevich 2006].

Dynamics
set of supported polychronous groups ≠ set of activated polychronous groups
Plan

1. Motivations
2. Problematics
3. Network architecture
4. Learning mechanisms
5. Results (1)
6. Polychronization
7. Results (2)
8. Conclusion
Figure: Activation ratio from 2000 to 5000 ms, and then from 8000 to 11000 ms.
Conclusion

- Algorithm easy to implement
- The learning seems to work on a classification task
- Easily explained by polychronization
- Activity easily monitored with polychronous groups
- Internal network is no longer a black-box contrary to ESN and LSM
Complex network analysis:
- Are polychronous groups the (or a part of the) link between topology and dynamics
- How far?
Thank you for listening.

Questions!

A supervised learning approach based on STDP and polychronization in spiking neuron networks
– Hélène Paugam-Moisy, Régis Martinez and Samy Bengio
Appendix

Plan

- Work in progress
- Reservoir Computing perspectives
- Groupes polychrones sur 100 neurones
- Modèle SRM0
- Modèle SRM1
- Forme d’un PPS
- Réseau expérimental
- Sensibilité à un motif spécifique
- Fenêtre STDP Eurich
- Fenêtre STDP classique
- Fenêtre STDP Meunier
- Stabilité du classifieur
- Codage temporel
- Architecture
- Activation des groupes polychrones
- Activité neuronale
Appendix

- Work in progress
- Reservoir Computing perspectives
- Groupes polychrones sur 100 neurones
- Modèle SRM0
- Modèle SRM1
- Forme d’un PPS
- Réseau expérimental
- Sensibilité à un motif spécifique
- Fenêtre STDP Eurich
- Fenêtre STDP classique
- Fenêtre STDP Meunier
- Stabilité du classifieur
- Codage temporel
- Architecture
- Activation des groupes polychrones
- Activité neuronale
- PG detection
- The model proposed
- Original problem
- Difference with synfire chain
- Network activity
Use larger inputs: encouraging tests with USPS dataset

2 versus 7: 96% success on train set, 93% on test set
3 versus 8: 89% success on train set, 86% on test set

- Switch to more than two classes
- Extend model with persistent activity
Might there be links with reservoir computing. Indeed, some theoretical properties exists : point-wise separation, universal approximation, echo state properties...
But still difficulties to investigate what’s going on in the reservoir (refering to special session)
Polychronous groups can be a reliable way
- to analyse dynamics of a spiking neuron reservoir
- to find optimal topologies (structures)
Groupes polychrones sur 100 neurones

[48] 18,24,80 (0,11,11) ==> 37 (16) — [49] 19,31,43 (3,0,5) ==> 6 (6)

[50] 19,55,76 (0,11,13) ==> 70 (16) — [51] 21,52,76 (7,7,0) ==> 11 (12)

> retour
Modèle $SRM_0$ (used)

\[ u_j(t) = \underbrace{\eta(t - t^f_j)}_{A: \text{refractory periode}} + \sum_i w_{ij} \underbrace{\epsilon(t - t^f_i - d_{ij})}_{B: \text{excitatory potential}} \]
Modèle $SRM_1$

\[ u_j(t) = \eta(t - t_{jf}) + \sum_i w_{ij} \sum_f \epsilon(t - t_{if} - d_{ij}) \]

A : refractory period
B : excitatory potential

\[ \eta(t) \]

\[ \epsilon(t) \]
Forme d’un PPS

\[ \text{exp}(-x/Tau) \]

![Graph showing the exponential decay of a PPS signal](image-url)
Réseau expérimental
Sensibilité à un motif spécifique

Neurones du sac

Neurone de sortie 1

retour
Fenêtre STDP Eurich

\[ W_\tau(x) \]

\[ W_\omega(x) \]
Fenêtre STDP classique

augmentation du poids

décalage temporel de la synapse (t\text{post} − t\text{pre} [ms])
Si $\Delta W \leq 0$, le poids est augmenté :

$$w_{ij} \leftarrow w_{ij} + \alpha \times (w_{ij} - w_{\text{min}}) \times \Delta W$$

Si $\Delta W \geq 0$, le poids est diminué :

$$w_{ij} \leftarrow w_{ij} + \alpha \times (w_{\text{max}} - w_{ij}) \times \Delta W$$
Stabilité du classifieur

Diagramme de Stabilité des réponses

Stabilité des réponses

Temps

Diagramme de Stabilité des réponses

Stabilité des réponses

Temps

retour
Codage temporel

Codage en intensité

Vecteur numérique

Vague de spikes dans un interval de codage temporel

Composantes du vecteur

Intensité

Neurones

Temps

Codage temporel
Architecture

K input cells

Internal network

M internal cells

2 output cells

input connections

test connection

output connections

class 1

class 2

with adaptable delay
Activation des groupes polychrones
Activité neuronale

Pre-synaptic neuron

dendrites

soma

axon

Post-synaptic neuron

synapse
To find all supported polychronous groups, we use the same algorithm as [Izhikevitch 2006]. It consists in scanning for spike time combination of all groups possible of 3 neurons (i.e. combinatorial questions), so that the spikes would trigger the firing of one or more impacted neurons, taking axonal delays into account.

Il est possible de procéder de même en cherchant plus de déclencheurs, mais la complexité est accrue: $O(n^p)$, avec $p$ nombre de déclencheurs.
The model proposed

Input layer (stimulation layer) :
- 10 neurons
- Outgoing connection probability : 0.1
- Delay to central assembly : 0 ms
The model proposed

Central assembly:
- 100 neurons, 80% excitators, 20% inhibitors
- Random topology
- Recurrent connection probability: 0.3
- Recurrent connections delay from 1 to 20 ms
- Spike Time Dependent Plasticity (STDP)
The model proposed

K input cells

Internal network

M internal cells

2 output cells

input connections

internal connections

output connections with adaptable delay

Output layer:

- 2 neurons: one for each target class
- Incoming connection probability: 1 (central assembly completely projected)
- Adaptable delays of input connections (all initialized to 10 ms)
Initial work

- Originally: problem for learning binary patterns
- Spike responses: all or nothing
- Solution: allow diversity in axonal delays

> when an initial neuron, A, fired, a second neuron, B, would fire 151ms later, followed by a third neuron, C, that would fire 289ms later with a precision across trials of 1 ms


> Synfire chains describe pools of neurons firing synchronously, not polychronously. Synfire activity relies on synaptic connections having equal delays or no delays at all. Though easy to implement, networks without delays are finite-dimensional and do not have rich dynamics to support persistent polychronous spiking.
Initialization phase
Learning phase observation

- Decreasing internal activity (STDP)
- Activity pattern different from an input to the other
- Margin evolution
Generalization performance

- Error rate with noise 4 : 4%
- Error rate with noise 8 : 19%