

# VORAIS: A MULTI-OBJECTIVE VORONOI DIAGRAM-BASED ARTIFICIAL IMMUNE SYSTEM

Luis Martí<sup>1,2</sup> Arsene Fansi-Tchango<sup>3</sup> Laurent Navarro<sup>3</sup> Marc Schoenauer<sup>1</sup>

<sup>1</sup> TAO team, INRIA/LRI(CNRS & UPSud), Université Paris-Saclay, Paris, France.

<sup>2</sup> Universidade Federal Fluminense, Niterói (RJ) Brazil.

<sup>3</sup> ThereSIS, Thales Group, Paris, France.



## Context

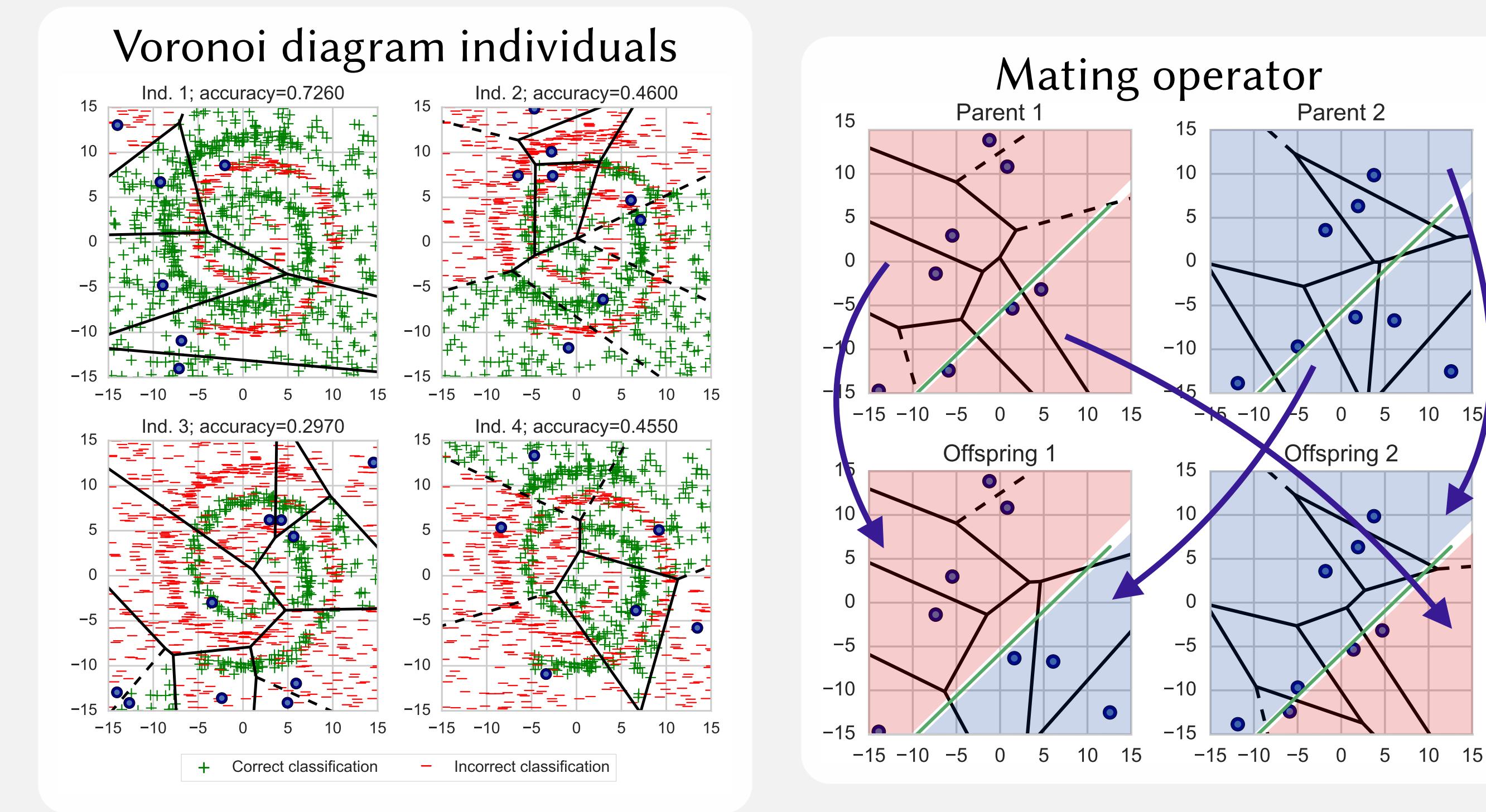
- ▷ Artificial Immune Systems have been derived from existing theories of the functioning of biological immune systems.
- ▷ They create a model that is able to discriminate between normal (self) and abnormal (non-self) objects.
- ▷ This feature makes AISs specially suited for dealing with problems related to anomaly and intrusion detection.

## Voronoi diagram-based Artificial Immune System

VorAIS models the self/non-self using a Voronoi diagram that determines which areas of the problem domain correspond to self or to non-self.

- ▷ Various classification metrics, like accuracy, recall, and specificity, must be taken into account.
- ▷ Multi-objective approach based on non-dominated sorting of NSGA-II.
- ▷ We introduce a mating operator as part of the AIS.

## VorAIS components



```
function mutate_voronoi( $\mathcal{I}, p_s, p_f, p_t, p_+, p_-, \eta$ )
     $\triangleright \mathcal{I}$ , individual to be mutated.
     $\triangleright p_s \in [0, 1]$ , prob. of mutating a site.
     $\triangleright p_f \in [0, 1]$ , prob. of mutating a site's feature.
     $\triangleright p_t \in [0, 1]$ , prob. of changing the label of a site.
     $\triangleright p_+ \in [0, 1]$ , prob. of adding a new site.
     $\triangleright p_- \in [0, 1]$ , prob. of removing a site.
     $\triangleright \eta \in (0, \infty)$ , learning rate.

    for all  $S \in \mathcal{I}$  do
        if  $U[0, 1] < p_s$  then
            for all  $x \in S$  do
                if  $U[0, 1] < p_f$  then
                     $x \leftarrow \text{mutate\_log\_normal}(x, \eta)$ 
                if  $U[0, 1] < p_t$  then
                     $S.\ell \leftarrow \text{switch\_label}(S.\ell)$ .
                if  $U[0, 1] < p_+$  then
                     $\mathcal{I} \leftarrow \mathcal{I} \cup \{\text{random\_site}\}$ .
                if  $U[0, 1] < p_-$  then
                     $i \leftarrow U[1, |\mathcal{I}|]$ .
                     $\mathcal{I} \leftarrow \mathcal{I} \setminus \{\mathcal{I}(i)\}$ .
    return  $\mathcal{I}$ , mutated individual.
```

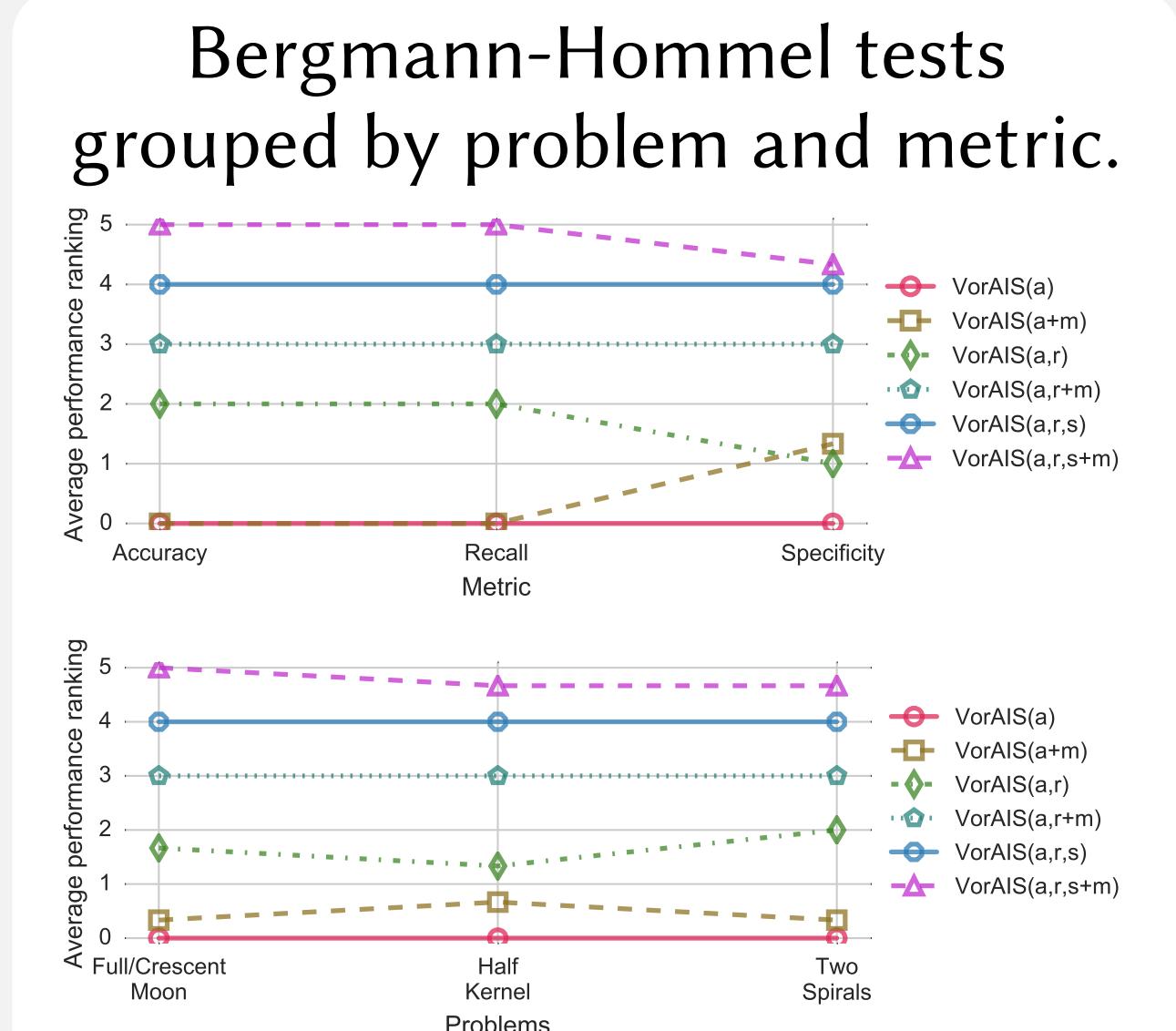
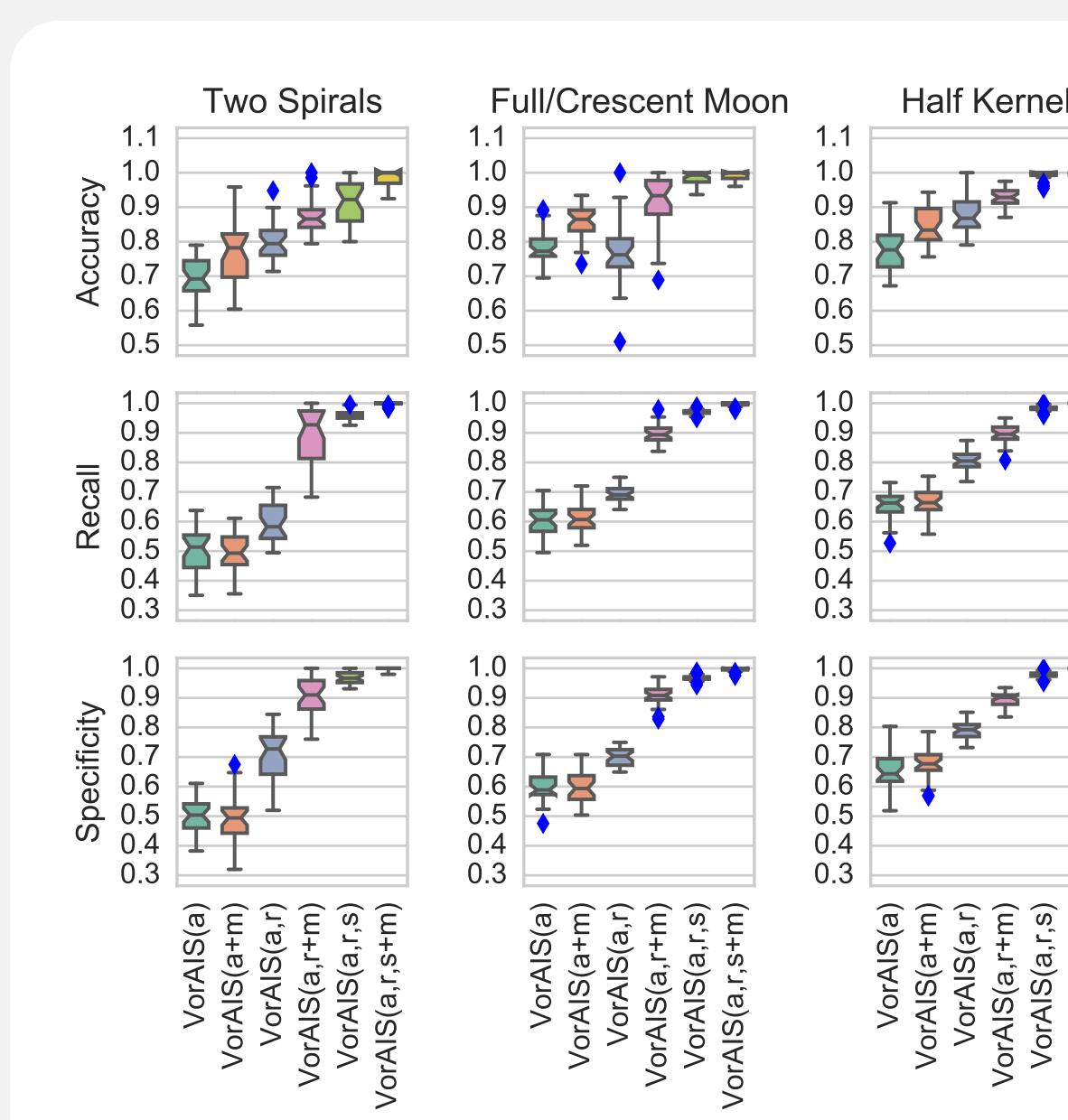
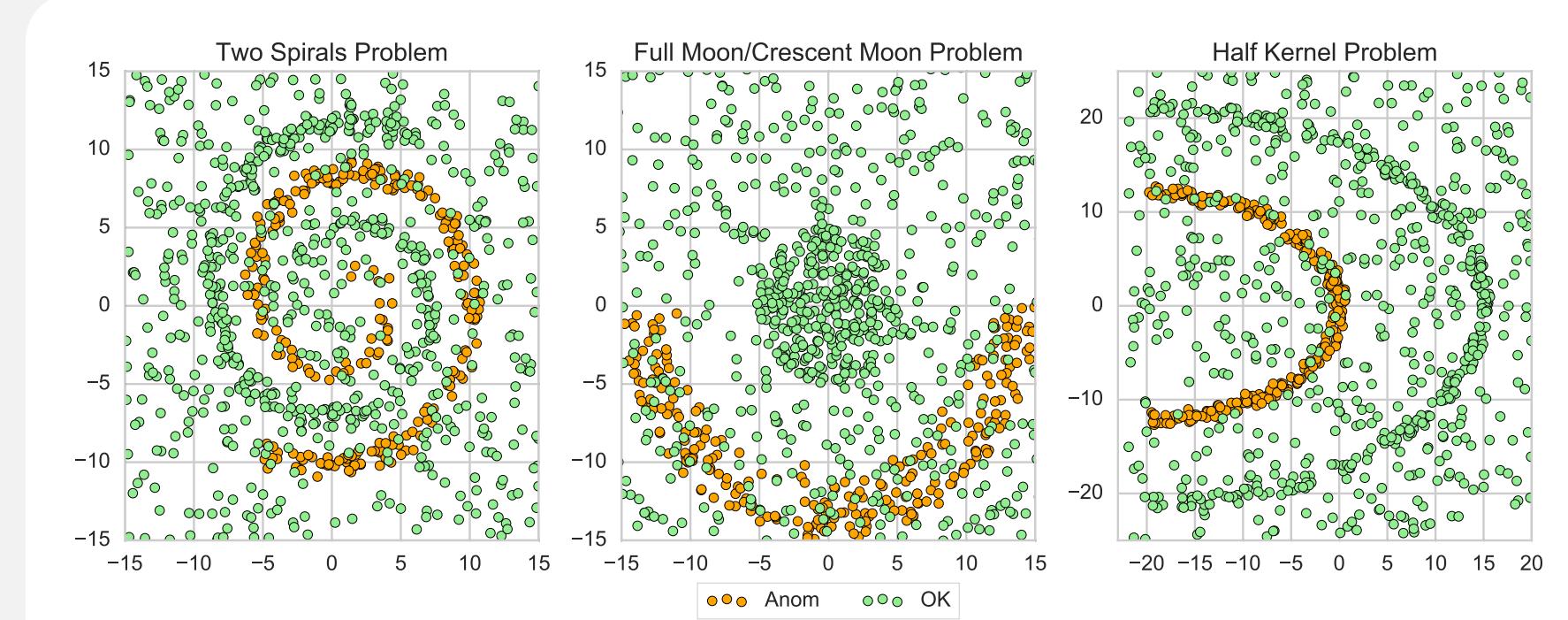
## VorAIS outline

1. Create an initial random population  $\mathcal{P}_0$  of  $n_{\text{pop}}$  individuals.
2. At iteration  $t$ , individuals in  $\mathcal{P}_t$  are mated and mutated.
3. An offspring population,  $\mathcal{P}_{\text{off}}$ , with  $n_{\text{off}}$  individuals is created.
4. From  $\mathcal{P}_t \cup \mathcal{P}_{\text{off}}$  the best  $n_{\text{pop}}$  individuals are selected using non-dominated sorting (with crowding distance).

Hyperparameters tuned by grid search (see PPSN paper).

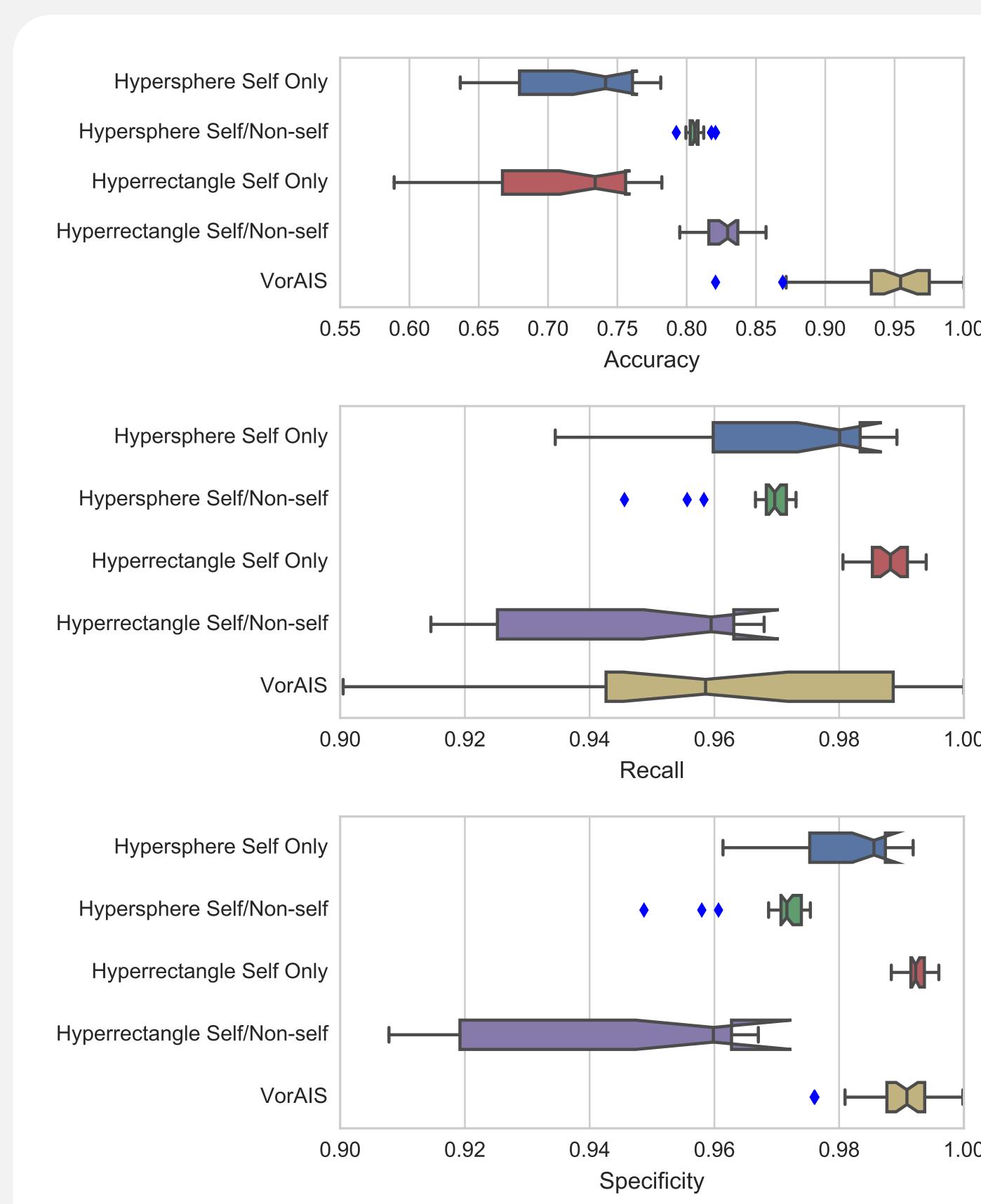
## Preliminary study on test problems

- ▷ Influence of the mating operator (+m) and the use of the classification metrics (a, r and s, respectively), on two-dimensional problems.



## NSL KDD'99 anomaly detection problem

- ▷ Compare VorAIS with other AISs in a computer network anomaly detection benchmark problem.
- ▷ NSL-KDD'99 has 41–38 numeric and 3 categorical—features and one class attribute describing the nature of the events.



	NSA <sub>sp</sub>	NSA <sub>sp</sub> <sup>+</sup>	NSA <sub>re</sub>	NSA <sub>re</sub> <sup>+</sup>	VorAIS
<b>Accuracy</b>					
NSA <sub>sp</sub>	×	—	~	—	—
NSA <sub>sp</sub> <sup>+</sup>	+	×	+	—	—
NSA <sub>re</sub>	~	—	×	—	—
NSA <sub>re</sub> <sup>+</sup>	+	+	+	×	—
VorAIS	+	+	+	+	×
<b>Recall</b>					
NSA <sub>sp</sub>	×	+	—	+	~
NSA <sub>sp</sub> <sup>+</sup>	—	×	—	+	~
NSA <sub>re</sub>	+	+	×	+	+
NSA <sub>re</sub> <sup>+</sup>	—	—	—	×	—
VorAIS	~	~	—	+	×
<b>Specificity</b>					
NSA <sub>sp</sub>	×	+	—	+	—
NSA <sub>sp</sub> <sup>+</sup>	—	×	—	+	—
NSA <sub>re</sub>	+	+	×	+	+
NSA <sub>re</sub> <sup>+</sup>	—	—	—	×	—
VorAIS	+	+	—	+	×

## Final remarks

- ▷ We have obtained a performance comparable with the state of the art but adequate classification performance is not enough.
- ▷ It is necessary to create a compact representation of the ‘normal’ data.
- ▷ We developed custom objective functions that rely on volume-based approaches.